

特開平10-320337

(43) 公開日 平成10年(1998)12月4日

(51) Int.Cl.<sup>6</sup>

G 0 6 F 13/00

識別記号

3 5 7

F I

G 0 6 F 13/00

3 5 7 Z

審査請求 未請求 請求項の数28 OL (全 24 頁)

(21) 出願番号 特願平9-125247

(22) 出願日 平成9年(1997)5月15日

(71) 出願人 000005108

株式会社日立製作所

東京都千代田区神田駿河台四丁目6番地

(72) 発明者 猪原 茂和

東京都国分寺市東恋ヶ窪一丁目280番地

株式会社日立製作所中央研究所内

(72) 発明者 鍵政 豊彦

東京都国分寺市東恋ヶ窪一丁目280番地

株式会社日立製作所中央研究所内

(72) 発明者 増岡 義政

東京都国分寺市東恋ヶ窪一丁目280番地

株式会社日立製作所中央研究所内

(74) 代理人 弁理士 小川 勝男

最終頁に続く

## (54) 【発明の名称】 分散データ管理方法

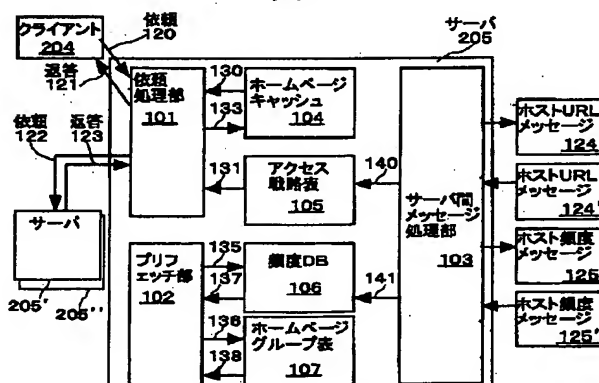
## (57) 【要約】

【課題】 インターネット使用人口の増加によるインターネットの不均一性と不安定性を低減し、ユーザに対してより快適な情報システムのサービスを行なうために、ユーザがクライアントに要求を発してから、その要求が満たされるまでの時間を短縮することが本発明の目的である。

【解決手段】 複数のサーバ(205'、205'')が全体で1つのサービスを提供する際、各サーバは過去の通信回線の状況(スループットやレイテンシ)を蓄積し、この蓄積情報をもとに任意のサーバ間で、通信回線の状況に応じたキャッシュ参照およびプリフェッチを行なう。

【効果】 インターネットの通信回線の不均一性と不安定性がある状況下でも、ユーザがクライアントに要求を発してから、その要求が満たされるまでの時間を短縮し、また特定の通信回線が不通でも、できる限りクライアントの依頼を処理することによりユーザにより快適な情報システムのサービスを行なう。

図1



【特許請求の範囲】

【請求項1】 1つ以上のプロセスを動作させる2つ以上の計算機がネットワークで接続されたコンピュータ・システム上で、1つ以上のデータを利用するクライアント対応のプロセスと、1つ以上のデータを保持し、クライアントからの依頼により指定されたデータを提供する2つ以上のサーバ対応のプロセスが用いる分散データ管理方法であって、

第1のサーバが必要とする必要データを1つ以上の第2のサーバが持つ可能性がある場合に、第1のサーバが、第1のサーバと第2のサーバが過去に行なった通信履歴を用いて第2のサーバから1つ以上の第3のサーバを選択して、第3のサーバに必要データの送信を依頼する分散データ管理方法。

【請求項2】 前記通信履歴として、前記第1のサーバと前記1つ以上の第2のサーバとの通信の、単位時間当りの通信可能データ量であるスループットと通信の遅延時間であるレイテンシを含む請求項1記載の分散データ管理方法。

【請求項3】 前記通信履歴として、前記必要データの大きさを含む請求項1記載の分散データ管理方法。

【請求項4】 前記通信履歴として、前記1つ以上の第2のサーバが前記必要データを保持しているかどうかを前記第1のサーバが予測したデータ保持確率を含む請求項1記載のデータ管理方法。

【請求項5】 前記必要データの通信完了時間を前記通信履歴から予測して、通信完了時間が短いことを基準として、前記1つ以上の第2のサーバから前記1つ以上の第3のサーバを選択する請求項1記載の分散データ管理方法。

【請求項6】 前記通信完了時間の予測を（前記レイテンシ+（前記必要データの大きさ/前記スループット））の式を用いて行なう請求項5記載の分散データ管理方法。

【請求項7】 前記データ保持確率と、前記第1の通信完了時間の積が小さいことを基準として、前記1つ以上の第2のサーバから前記1つ以上の第3のサーバを選択する請求項4記載の分散データ管理方法。

【請求項8】 2つ以上のサーバから通信が行なわれた際に、前記通信履歴中の該2つ以上のサーバの通信履歴の両方を更新する請求項1記載の分散データ管理方法。

【請求項9】 1つ以上のプロセスを動作させる2つ以上の計算機がネットワークで接続されたコンピュータ・システム上で、2つ以上のサーバ対応のプロセスが用いる分散データ管理方法であって、第1のサーバの必要データを2つ以上の第2のサーバが持つ可能性があり、第1のサーバが該2つ以上の第2のサーバに対して必要データの送信を依頼する場合に、第1のサーバが、該2つ以上の第2のサーバから1つ以上の第3のサーバを選択し、該第3のサーバには前記必要データを直ちに送信す

ることを依頼し、第3のサーバ以外の第2のサーバには前記必要データの送信を待機することを依頼する分散データ管理方法。

【請求項10】 第1のサーバが、第1のサーバと第2のサーバが過去に行なった通信履歴を用いて、前記第2のサーバから前記第3のサーバを選択する請求項9記載の分散データ管理方法。

【請求項11】 前記第3のサーバの一部または全部から、前記必要データが送信できない旨が前記第1のサーバに対して通信された際に、前記第1のサーバが、前記第3のサーバに含まれない前記第2のサーバの一部または全部に対して、前記必要データを直ちに送信することを依頼する請求項10記載の分散データ管理方法。

【請求項12】 前記第2のサーバのうち2つ以上のサーバから、前記必要データの送信が開始された際に、当該2つ以上のサーバのうち1つ以上の第4のサーバからの送信のみを継続し、他のサーバからの送信を中止することにより、不要な必要データの送信を抑制する請求項11記載の分散データ管理方法。

【請求項13】 前記通信履歴を用いて、前記第3のサーバの選択と第4のサーバの選択を行なう請求項12記載の分散データ管理方法。

【請求項14】 1つ以上のプロセスを動作させる2つ以上の計算機がネットワークで接続されたコンピュータ・システム上で、1つ以上のクライアント対応のプロセスと2つ以上のサーバ対応のプロセスが用いる分散データ管理方法であって、第1のサーバが、該第1のサーバと1つ以上の第2のサーバとの過去の通信に関する、第1の時間にわたる第2の時間間隔毎の通信の履歴である時間毎通信履歴を持つ分散データ管理方法。

【請求項15】 前記時間毎通信履歴として、前記第1のサーバと前記第2のサーバの間の、第1の時間にわたる第2の時間間隔毎の通信のスループットの履歴と、第1の時間にわたる第2の時間間隔毎のレイテンシの履歴を含む請求項14記載の分散データ管理方法。

【請求項16】 第1のサーバが、今後クライアントから依頼される可能性が高いと判断されるプリフェッチ対象データをクライアントからの依頼以前に他のサーバに依頼してプリフェッチする際に、プリフェッチ対象データを持つ可能性がある1つ以上の第2のサーバが存在する場合、前記時間毎通信履歴を用いて当該第2のサーバからプリフェッチ対象データが高速に入手可能な時刻を予測して、当該時刻に第1のサーバが、第2のサーバの一部または全部である第3のサーバに対してプリフェッチ対象データの送信を依頼する請求項14記載の分散データ管理方法。

【請求項17】 1つ以上のプロセスを動作させる2つ以上の計算機がネットワークで接続されたコンピュータ・システム上で、2つ以上のサーバ対応のプロセスが用いる分散データ管理方法であって、第1のサーバが、第1

のサーバと第2のサーバが過去に行った通信履歴を用いて第1のサーバが知る第2のサーバから1つ以上の第3のサーバを選択し、第1のサーバが持つデータの一覧の一部または全部を、第3のサーバに通信する分散データ管理方法。

【請求項18】前記一覧の通信が行なわれた時刻と現在時刻の差から、前記データ保持確率を決定する請求項17記載の分散データ管理方法。

【請求項19】前記一覧の通信が行なわれた時刻に、前記データ保持確率を1とし、以後一定時刻毎に、前記データ保持確率を減少させる請求項18記載の分散データ管理方法。

【請求項20】1つ以上のプロセスを動作させる2つ以上の計算機がネットワークで接続されたコンピュータ・システム上で、1つ以上のクライアント対応のプロセスと2つ以上のサーバ対応のプロセスが用いる分散データ管理方法であって、第1のサーバが、クライアントからの依頼の履歴から、第1のデータの依頼に続けて依頼される頻度が高かった1つ以上の第2のデータを決定して、該第1のデータの名前と該1つ以上の第2のデータの名前を組にしたデータを参照関係情報として保持するとともに、第1のサーバの参照関係情報と、第1のサーバと同じく参照関係情報を保持する1つ以上の第2のサーバの参照関係情報とを交換する分散データ管理方法。

【請求項21】前記第1のサーバが、クライアントから第1のデータの依頼を受けた際に、前記参照関係情報を用いて当該第1のデータに続いて依頼される可能性が高い1つ以上の第2のデータを決定し、当該第2のサーバを持つ可能性のある1つ以上の第2のサーバに対して第2のデータのプリフェッチを行なう請求項20記載の分散データ管理方法。

【請求項22】前記1つ以上の第2のサーバを、前記第1のサーバが知る1つ以上の第3のサーバから、第1のサーバと第2のサーバが過去に行った通信履歴を用いて選択する請求項21記載の分散データ管理方法。

【請求項23】1つ以上のプロセスを動作させる2つ以上の計算機がネットワークで接続されたコンピュータ・システム上で、1つ以上のクライアント対応のプロセスと2つ以上のサーバ対応のプロセスが用いる分散データ管理方法であって、

第1のサーバが、他のサーバから第1のデータの依頼を受けた際に、該第1のデータに続いて依頼される可能性の高い第2のデータがある場合に、該第2のデータを持つ可能性のある1つ以上の第2のサーバに対して、該第2のデータの送信を依頼することにより、プリフェッチをあるサーバから1つ以上の他のサーバへ階層的に行なう分散データ管理方法。

【請求項24】前記1つ以上の第2のサーバを、前記第1のサーバが知る1つ以上の第3のサーバから、第1のサーバと第2のサーバが過去に行った通信履歴を用いて

選択する請求項23記載の分散データ管理方法。

【請求項25】1つ以上のプロセスを動作させる2つ以上の計算機が、ネットワークで接続されたコンピュータ・システム上で、1つ以上のクライアント対応のプロセスと2つ以上のサーバ対応のプロセスが用いる分散データ管理方法であって、

第1のサーバの保持している2つ以上のデータの一部を破棄する際に、該データを他のサーバから得るのに必要とする時間の予測を用いて、該2つ以上のデータを順位づけする分散データ管理方法。

【請求項26】前記第1のサーバの保持している2つ以上のデータの一部を破棄する際に、該データを他のサーバから得るのに必要とする時間の予測と、該データが過去の一定時間内にクライアントから依頼された回数とを用いて、該2つ以上のデータを順位づけする請求項25記載の分散データ管理方法。

【請求項27】前記データのそれぞれを他のサーバから得るのに必要とする時間の予測を、サーバ間で過去に行った通信履歴を用いて行なう請求項25記載の分散データ管理方法。

【請求項28】1つ以上のサーバ対応のプロセスを動作させる2つ以上の計算機がネットワークで接続されたコンピュータ・システム上で、第1のサーバの必要データを1つ以上の第2のサーバが持つ可能性がある場合に、第1のサーバが、第2のサーバを含む複数のサーバに必要データの送信を依頼する分散データ管理方法であって、

第2のサーバが、他のサーバの依頼を一定時間内に一定回数以下のみ受け付けるか、または依頼に対するデータの送信を一定時間内に一定量以下のみ行なうことにより第2のサーバが過負荷になることを避ける分散データ管理方法。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】本発明は、計算機システムに関し、例えばネットワークによって接続された複数の計算機が、情報を流通・共有・変更するシステム（情報システム）内の情報を管理する分散データ管理方法に関し、特にWorld Wide Web (WWW) に好適な分散データ管理方法に関する。

【0002】

【従来の技術】以下の記述に用いるいくつかの用語をまず説明する。

【0003】インターネット上のWWW、匿名FTPをはじめとする情報システムは、通常分散コンピュータ・システムの一形態である「クライアント・サーバ・システム」の形態で構築されている。クライアント・サーバ・システムでは、システム全体の処理は2種類に分割され、第1の部分を「サーバ」と呼ばれる実行中のプログラム（以下、プロセスという）で処理し、第2の部分を

複数の「クライアント」と呼ばれるプロセスで処理する。情報システムのクライアントは通常、個人または企業内のユーザが操作するコンピュータで動作する。情報システムのサーバは、クライアントに提供する情報を格納する。情報システムのクライアントは、サーバに新たな情報を格納したり、サーバに要請して情報を取得する。

【0004】コンピュータ・システムでは、1つの情報を複数の場所に一時的にコピーしておくことにより、当該情報へのアクセスを高速にしたり、アクセスできる可能性を高めることが広く行なわれる。このようなコピーはヒント、キャッシュ、レプリカ、スタッシュ等と呼び分けられる(Sape Mullender編の文献" Distributed Systems (1st ed.)" pp.13-15, ACM press, 1989参照)。以下ではこれらのコピーを総称して「キャッシュ」と呼ぶ。またキャッシュを作ることを「キャッシュする」と記す。

【0005】WWWでは、WWWサーバが、発信する情報を「ホームページ」と呼ばれる単位で保持している。ホームページは、それぞれURL (Uniform Resource Locatorの略) と呼ばれる名前を持っている。URLはWWWで用いられるプロトコルと、情報源であるコンピュータのホスト名、および情報源において特定のデータを指定することが可能な文字列である。"http://www.hitachi.co.jp/index.html" は URLの一例である。

【0006】なお、一般にURLは(ホームページのデータや画像データ等の)ひとかたまりのデータに対応づけられるが、以後このひとかたまりのデータのことを「URLに対応する情報」または「URL内容」と記す。また、第1のURLに対応する情報中に含まれる第2のURLは「ハイパー・テキスト・リンク(または簡単にリンク)」とよばれる。第1のURLに対応する情報中に第2のURLが含まれることを、以下「第1のURLから第2のURLにリンクがある」と記す。

【0007】WWWで用いられている技術(公知例1)を説明する。

【0008】WWWクライアントのユーザは、アクセスしたいホームページのURLをWWWクライアントに指定する。WWWのサーバとクライアントの第1の処理形態では、WWWクライアントは当該URLの第2要素で指定されるWWWサーバに対して、当該URLに対応するホームページの送信を依頼する。当該WWWサーバは、この要求に応じて、当該ホームページを当該WWWクライアントに与える。

【0009】また第2の処理形態では、WWWクライアントが、ユーザが与えたURLの第2要素で指定されるWWWサーバに依頼を行なうかわりに、「プロキシ・サ

ーバ」と呼ばれる第2のサーバに依頼を行なう場合もある。当該第2のサーバは、当該WWWサーバから当該URLに対応するホームページを得るか、さらに別のプロキシ・サーバに当該URLに対応する情報取得を依頼する。このプロキシ・サーバの依頼が複数段階である場合、プロキシ・サーバは親子関係を持つ。親子関係を持つプロキシ・サーバに関しては、例えばA.Chankhunthod他の文献"A Hierarchical Internet Object Cache"(1996 USENIX Technical Conference, pp.153-163, 1996)に記載されている。

【0010】なお、WWWのクライアントやプロキシ・サーバは、キャッシュを持つことができる。クライアントが持つキャッシュは、以前に当該クライアントが取得したホームページを保持しており、当該クライアントのみが利用する。プロキシ・サーバがもつキャッシュは、当該プロキシ・サーバが1つ以上のクライアントまたは他の1つ以上のプロキシ・サーバまたはその両方の依頼によって取得したホームページを保持しており、当該プロキシ・サーバを利用するクライアントおよびプロキシ・サーバによって共用される。

【0011】ネットワーク・ニュース・システム(以後公知例2と呼ぶ。B.Kantor他著の文献"Network News Transfer Protocol: A Proposed Standard for the Stream-Based Transmission of News" Network Working Group RFC-977等に記載されている)は、1つ以上のサーバによって構成される。ユーザは通常、クライアントを用いて当該複数のサーバのうち1つを選択して利用する。ネットワーク・ニュース・システムにおける情報の単位は「ニュース記事」と呼ばれる。ユーザは通常、クライアントを用いてニュース記事をサーバに提供する操作と、ニュース記事をサーバから取り出す操作を行なう。ユーザがニュース記事が第1のサーバに提供すると、第1のサーバは1つ以上の第2のサーバに当該ニュース記事のコピーを送り、さらに第2のサーバが別のサーバに当該ニュース記事のコピーを送る、という具合で、最終的には上記複数のサーバすべてに当該ニュース記事がコピーされる。

【0012】公知例3として、広域名前サービスDomain Name Systemを説明する。(以下DNSと略す。例えばP.Mockapetris著の文献"Domain Names—Implementation and Specification" Network Working Group RFC-1035の特に第2節に記載されている)DNSは主にシンボリックなホスト名と、そのホスト名の付随情報(IPアドレスやメールサーバ)とを対応づける。複数のD

NSサーバは木構造を構成する。クライアントの依頼は当該木構造をたどって、複数のサーバの間を転送され、処理される。DNSのクライアントであるresolverは、あるホスト名に対応するホスト関連情報をサーバの1つに依頼する。この際当該サーバは、当該ホスト名に対応するホスト関連情報をクライアントに返すか、当該DNSサーバの親サーバ(DNSサーバの木構造において根に一段近いDNSサーバ)に当該依頼を転送する。なお親サーバは、直接の子サーバ群がどのホスト関連情報を持つかを把握している。このため、木構造の根まで依頼が転送されると、今度は木構造を下って依頼を処理できるDNSサーバの方に依頼を転送する。依頼は最終的に、ホスト関連情報を保持するDNSサーバに到達して、クライアントにホスト関連情報が返されるか、木構造を下っていく段階で、当該依頼のホスト名に対応するホスト関連情報がどのDNSサーバからも提供されていないことが判明して、クライアントに失敗が返される。

【0013】ローカル・エリア・ネットワーク(LAN)内の分散ファイルシステムにおいて、キャッシュのための空間を複数のコンピュータで共有する方法も知られている(公知例4)。Michael Dahlin他著の文献"Cooperative Caching: Using Remote Client Memory to Improve File System Performance"(First USENIX Symposium on Operating Systems Design and Implementation, pp. 267-280, 1994)では、クライアントはまず「マネージャ」と呼ばれるサーバに、ファイルブロックを依頼する。マネージャは、どのファイル・ブロックがどのコンピュータに格納されているかを保持しており、クライアントにファイル・ブロックが格納されているコンピュータを返答するか、当該コンピュータにクライアントの要求を転送する。同様の方法としては、M. J. Feeley他著の文献"Implementing Global Memory Management in a Workstation Cluster"(ACM 15th Symposium on Operating Systems Principles, pp. 201-212, 1995)や P. Sarkar他著の文献"Efficient Cooperative Caching Using Hints"(Second USENIX Symposium on Operating Systems Design and Implementation, pp. 35-46, 1996)が知られている。マネージャは複数存在することができるが、ファイル・ブロックとマネージャとの対応関係はあらかじめ定められており、全クライアントおよび

サーバがその対応関係を知っている。この対応関係はシステム動作中に変化しない。

【0014】公知例5および公知例6として、cache-coherent non-uniform memory access (CC-NUMA) や cache only memory access (COMA) と呼ばれる種類の計算機で用いられている技術を説明する。CC-NUMA計算機やCOMA計算機では、メモリの断片(キャッシュライン)を多数のプロセッサの近くにキャッシュし、それらキャッシュの間の一貫性を保つ機構が含まれている。特に以下の2つの方法が知られている。

【0015】第1の方法(公知例5)は、上記マネージャに相当する「ホーム」と呼ばれる処理装置またはデータが、メモリの断片がどのプロセッサによってキャッシュされているかを把握しておく方法である。Jeffrey Kuskina他著の文献"The Stanford FLASH Multiprocessor"(Proceedings of the 21st Annual International Symposium on Computer Architecture, pp. 302-314, ACM, 1994)に記載のシステム等で第1の方法が用いられている。

【0016】第2の方法(公知例6)は、キャッシュの作成と削除と通信方法に一定の制約を課すことによって、一定回数の通信(通常マルチキャストまたはブロードキャストを含む)でキャッシュの特定と一貫性の保持ができることを保証する方法である。Henk L. Muller他著の文献"The Data Diffusion Machine with a Scalable Point-to-Point Network"(Technical Report CSTR-93-17, Department of Computer Science, University of Bristol, October 1993)に記載のシステム等で第2の方法が用いられている。

【0017】

【発明が解決しようとする課題】現在のインターネット全体の通信性能はユーザの要求する速度に遠く及んでおらず、多くの不安定要因を抱えている。WWWの爆発的な普及により、多くのワイド・エリア・ネットワーク(WAN)で混雑が起こっている。高速なバックボーン通信回線が日々増強されている一方で、家庭からのインターネット利用者は、一般的なLANよりも非常に遅い通信回線を用いてインターネットに接続している。現在でもWWWサーバとインターネットのユーザ数は増加の一途をたどっており、1996年1月のある統計では、全世界のインターネット利用のコンピュータは900万台、6ヶ月弱で2倍の速度で増加している。

【0018】このため、インターネットの通信回線には不均一性と不安定性がある。ここでの不均一性とは、様々な通信回線が混在していることを指す。例えば、通信回線によってスループット（単位時間当りの通信データ量）とレイテンシ（通信が受ける遅延時間）が異なることがあげられる。ここでの不安定性とは、通信回線のスループットとレイテンシが刻一刻と変化し、場合によっては通信不能になったりすることを指す。例えば、時間帯や曜日によって通信回線の混雑が変化したり、ある通信回線の増強によってルーティングのパターンが変更され、別の通信回線が混雑したり混雑が解消されたりすることを指す。

【0019】このような状況を鑑み、ユーザにより快適な情報システムのサービスを行なうために、ユーザがクライアントに要求を発してから、その要求が満たされるまで時間を短縮することが本発明の目的である。以下この目的を実現する上での課題と、公知の技術との隔たりを説明する。

【0020】第1に、公知の技術では、ある通信回線が不安定な状況下では、たとえ別の通信回線が正常に動作していてもクライアントの依頼を高速に処理できない場合があった。

【0021】公知例1のWWWのクライアントやWWWのプロキシ・サーバは、URLによって指定された特定のWWWサーバまたは特定のプロキシ・サーバと通信する。このため、当該サーバ（またはプロキシ・サーバ）への通信回線が混雑していたり、不通である場合には、たとえ別の通信回線が正常に動作していても、クライアント（またはプロキシ・サーバ）がホームページの入手に長時間を要したり、ホームページの入手ができない。また同じ理由で、ある時点で低速だった通信回線が増強等により高速になったとしても、その利益を享受することが困難だった。公知例2から公知例5でも、クライアントとサーバ、またはサーバと別のサーバの通信が、あらかじめ定められた相手とのみ行なわれる点は公知例1と共通しており、上記の問題がある。一方公知例6は、1つの計算機または物理的に密接に結び付いた複数の計算機で使用する技術であり、マルチキャストまたはブロードキャストが不具合なく動作することを前提としている。このためLAN環境やWAN環境にひろく適用することは困難である。また同じ理由で、特定の通信回線が混雑していたり、不通である場合には（たとえ別の通信回線が正常に動作していても）、キャッシュラインの入手に長時間を要したり、キャッシュラインの入手ができない。

【0022】第2に、公知の技術では通信回線が不安定な状況下でキャッシュの利用度を上げることが達成されていなかった。

【0023】公知例1、2、3、5では、複数のサーバ（またはクライアントとサーバ）がキャッシュを持つ場

合、これら複数のキャッシュは階層構造をなす。クライアントが発する依頼には、当該依頼を処理する責任を持つ特定の第1のサーバが存在している。当該依頼は、当該第1のサーバに直接送信されるか、当該第1のサーバとクライアントの間の1つ以上の第2のサーバの間に順に転送される（この場合、当該1つ以上の第2のサーバは依頼の内容によって一意に決定される）。前者の場合、当該依頼は当該第1のサーバのキャッシュしか利用できない。また後者の場合には、当該1つ以上の第2のサーバのキャッシュしか利用できない。すなわち第1、第2のサーバの他のサーバのキャッシュは利用されず、キャッシュの利用度が低い。例えば公知例1のWWWのプロキシ・サーバは親子関係を持つので、プロキシ・サーバA、B、Cが存在してAが親、BとCがAの子として動作している場合、BはAに依頼を転送することによってAのキャッシュを利用することができるが、BがCのキャッシュを利用することはできない。また公知例2では、各ニュース記事は、当該ニュース記事を必要とするすべてのサーバにコピーされる。このため、各サーバは他のサーバのキャッシュを全く利用せず、したがって各サーバが大規模な二次記憶を用意する必要がある。一方、公知例4はLAN環境を仮定しており、キャッシュの利用度を上げる目的で、1つのデータのコピーが複数のキャッシュ中でできるだけ少数になるような制御が行なう。しかし、通信回線が不安定な場合には、クライアントの依頼が少数のキャッシュに到達できない場合があり、逆にキャッシュの利用度が低くなる恐れがある。公知例6は、複数のキャッシュのうち最も短時間で利用できるキャッシュを選択するが、この際にマルチキャストまたはブロードキャストが用いられる。このため、公知例6の方法を通信回線が不安定になる可能性があるLAN環境やWAN環境にひろく適用することは困難である。

【0024】第3に、公知の技術ではデータとデータの間の参照関係の使用頻度を複数のサーバが交換する方法を持たなかった。

【0025】WWWのホームページとハイパー・テキスト・リンクに代表されるように、情報システムが提供するデータには参照関係があり、参照関係のなかでも頻繁にたどられる参照関係（使用頻度の高い参照関係）があることが多い。参照関係の使用頻度情報は、プリフェッチなどに利用できる。頻繁にたどられる参照関係で結ばれた一連の情報をクライアントからの依頼以前にプリフェッチしておけば、クライアントからの依頼があった時に通信回線が不安定でも、当該一連の情報の依頼は高速に処理することができるためである。特に複数のサーバが存在する場合、参照関係の使用頻度情報は、個々のサーバがそれぞれ収集するより、複数のサーバが収集した情報を交換して集計する方がより確からしい情報となる。しかし公知例1から6では、参照関係の使用頻度情



報を複数のサーバが交換する方法を持たなかった。このため、参照関係の使用頻度情報の確からしさに限界があり、プリフェッチの効果も限られていた。

【0026】第4に、公知の技術ではキャッシュの入れ替えに通信回線の不均一性と不安定性が考慮されていなかったため、重要なキャッシュを破棄する恐れがあった。

【0027】公知例1、2、3では、各サーバ（またはクライアント）が、利用履歴等に基づいてそれぞれのキャッシュの入れ替えを行なう。公知例4はLAN環境での使用を想定しているため、キャッシュの入れ替えの際のキャッシュの優先度づけに、通信回線の状態は考慮されない。また、公知例5と6も、1つの計算機または物理的に密接に結び付いた複数の計算機で使用することを想定しているため、キャッシュの入れ替えの際のキャッシュの優先度づけに、通信回線の状態は考慮されない。この結果、公知例1から6のいずれでも、現在の通信状態では再び入手することができない、または入手に長時間を要するキャッシュを破棄する場合がある。

【0028】第5に、公知の技術ではいずれも、第1のサーバが1つ以上の第2のサーバからの依頼を受ける場合に、第2のサーバが第1のサーバを無制限に利用し、結果として第1のサーバを過負荷におとしいることに對する対策がなかった。このため、複数のサーバが複数のユーザの組織を越えて依頼を行ないあうことが困難だった。公知例1、2、3は、クライアント・サーバ・システムであるため、サーバが依頼を拒否すると、クライアントの依頼が処理されない。一方、公知例4、5、6は限られた範囲の複数のサーバが依頼を行ないあうため、第1のサーバの過負荷を問題にする必要がなかった。

【0029】

【課題を解決するための手段】本発明では、複数のサーバを用い、複数のキャッシュを複数のサーバが保持することによって、1つの情報システムのサービスを提供する。

【0030】第1の課題を解決するため、以下の3つの手段を講じる。

【0031】(1) 第1のサーバが必要とするデータ（以後必要データと呼ぶ）を入手しようとする際、当該必要データは本発明の方法で用いる複数のサーバのうち、2つ以上の第2のサーバにキャッシュされている可能性がある。この際第1のサーバは、第1のサーバと第2のサーバが過去に行なった通信履歴を保持し、当該通信履歴を用いて第2のサーバから1つ以上の第3のサーバを選択して、第3のサーバに必要データの送信を依頼する。この通信履歴から第1のサーバは、現時点で高速に必要データの入手が可能と予測される第3のサーバを選択できる。このため、高速に通信できる通信回線を選択的に用いてクライアントの依頼を処理可能となる。

【0032】(2) 第1のサーバは、必要データを入手するための第3のサーバを2つ以上選択し、これら複数のサーバに対して同時に依頼を送信することにより、クライアントの依頼を高速に処理することを試みる。これだけでは当該複数のサーバからほぼ同時に依頼の返答が開始されて、却って通信の速度を下げる恐れがある。そこで、当該2つ以上の第2のサーバのうち、1つ以上の第3のサーバには必要データを直ちに送信することを依頼し、別の1つ以上の第4のサーバには、必要データの送信を待機することを依頼する。第3のサーバからの必要データの送信が長時間を要するか、不可能であることが判明した時点で、すでに待機している第4のサーバから直ちに必要データの送信を受けられる。このため、通信回線の状況が変化した場合にもクライアントの依頼を高速に処理可能となる。

【0033】(3) 第1のサーバは、当該第1のサーバと1つ以上の他のサーバとの過去の通信に関する、第1の時間にわたる、第2の時間間隔毎の通信の履歴（時間毎通信履歴）を持つ。当該時間毎通信履歴は、通信回線の状況が周期的に変化する場合（例えば昼間は混雑しているが夜間は混雑していない通信回線を用いる場合等）に、第1のサーバが通信を行なう時刻の選択を可能にする。これら3つの手段によって本発明は、高速に通信できる通信回線を選択的に利用し、若しくは高速に通信できる時間帯に選択的に利用して、通信回線が不安定または不均一な状況下で、クライアントの依頼を高速に処理する。

【0034】第2の課題を解決するために、第1のサーバは、当該第1のサーバが持つキャッシュの一覧の一部または全部を、1つ以上の第2のサーバに通信する。この際当該1つ以上の他のサーバは前記通信回線の状況の履歴を用いて選択する。これにより、第1のサーバのキャッシュの一覧を、その時点で高速に通信できる第2のサーバに通信することができる。このため、第1のサーバと第2のサーバの間の通信回線が正常な場合には、他の通信回線が不安定でも、第2のサーバが第1のサーバのキャッシュを利用可能となる。

【0035】第3の課題を解決するために、クライアントが第1のデータを依頼した後に、1つ以上の第2のデータを依頼する可能性が高い場合に、第1のサーバは第1のデータと第2のデータの対応関係（参照関係情報）を持ち、かつ当該参照関係情報を1つ以上の第2のサーバに送信する。この手段を用いて本発明の複数のサーバは、参照関係情報を他のサーバと交換し、頻繁にたどられる参照関係の使用頻度の確からしさを向上させる。これにより頻繁に依頼される一連のデータをクライアントに高速に提供する。

【0036】また参照関係情報をプリフェッチに用いることによって、多数のデータを同時に依頼することが可能になる。この際本発明ではサーバが複数あることを利

用して、複数のサーバが階層的なパイプライン状にプリフェッチを行ない、特定のサーバにボトルネックを生じたり、特定のネットワークが過負荷になってプリフェッチが効果的に行なわれない等の問題を避ける。

【0037】第4の課題を解決するために、第1のサーバは前記通信回線の状況の履歴を用いて情報の単位のアクセス頻度および最終アクセス時間、平均スループット、平均レイテンシ、他のサーバのキャッシュの内容から、キャッシュの優先度づけを行ない、この優先度に基づいてキャッシュの入れ替えを行なう。これにより、現在の通信状態では再び入手ができない、または入手に長時間を要するキャッシュを破棄することを避けることが可能となる。

【0038】第5の課題を解決するために、第1のサーバが、1つ以上の第2のサーバの必要データの送信の全体を一定時間内に一定量以下、または一定時間内に一定回数以下に制限する手段を持ち、一方で第2のサーバの各々は、第1のサーバを含む2つ以上の第3のサーバに、必要データの送信を依頼する。これにより、第1のサーバが第2のサーバからの依頼によって過負荷になることを防ぎつつ、第2のサーバの処理を妨げないことが可能となる。本発明が複数のサーバからなっており、ひとつのサーバが依頼を拒否しても、他のサーバが依頼を処理することが可能なため、この手段が有効となる。

【0039】これらの手段によって、本発明は、インターネットの通信回線の不均一性と不安定性がある状況下でも、ユーザがクライアントに要求を発してから、その要求が満たされるまで時間を短縮する。また特定の通信回線が不通でも、できる限りクライアントの依頼を処理する。これによりユーザにより快適な情報システムのサービスを行なう。

#### 【0040】

【発明の実施の形態】以下、本発明の実施の一形態を、図面を参照しながら説明する。なお、簡単のため、本明細書中では「発明の実施の形態」を、単に「実施例」と呼ぶことにする。

#### 【0041】全体構成

本実施例は、他の情報システムから入手する情報とユーザが入力する情報を、分散した1つ以上のキャッシュに保持する分散データ管理方法を搭載した計算機システムである。図2に本実施例の全体を示す。本実施例の全体201は分散コンピュータ・システムであり、ネットワーク202と、ネットワーク202によって相互接続された1つ以上の計算機203、203'、203''、…からなる。

【0042】ネットワーク202は、ある団体（企業や学校や類似の団体）の全体や位置部門でよく使用されるLANでもよく、また地理的に分散した複数の地点を結合するWANの一部または全部でもよい。またネットワーク202は、計算機間結合網や並列計算機内部のプロ

セッサ要素間の結合網でもよい。

【0043】計算機203、203'、203''、…は、それぞれ1個以上のクライアントまたは1個以上のサーバまたはその両方を動作させる。本実施例の全体201で、少なくとも1つのサーバ205と少なくとも1つのクライアント204が存在する。なお以下で、本実施例の全体に含まれるサーバ群を「本システム」と呼ぶことがある。計算機203、203'、203''、…は、いわゆるパーソナル・コンピュータ、ワークステーション、並列計算機、大型計算機等、任意のコンピュータでよい。また、クライアント204、204'、204''、…を動作させる計算機203、203'、203''、…は、サーバ205、205'、205''、…と通信する機能をもっていればよい。すなわち各種コンピュータ端末装置、携帯通信端末（いわゆるパーソナル・デジタル・アシスタンスPDAやハンドヘルド・パーソナル・コンピュータHPC）、ネットワーク・コンピュータなどでも差し支えない。なお、図2に示した計算機203、203'、203''、…、クライアント204、204'、204''、…、およびサーバ205、205'、205''、…の数と構成は、例として示したもので、本発明の範囲を限定するものではない。

【0044】クライアント204は、WWWで使われる各種プロトコル、すなわちHTTP（Hypertext Transfer Protocolの略）プロトコル、FTP（File Transfer Protocolの略）プロトコル等を用いてサーバ205と通信を行なう。クライアント204は、グラフィック・ユーザ・インターフェースを備えたWWWブラウザ、テキスト表示のみを行なうWWWブラウザ、ワード・プロセッサや表計算ソフトウェアに内蔵されたWWWのクライアント、WWWのプロキシ・サーバ等でよい。なお、クライアント204が用いるプロトコル、およびクライアント204の種類は例として示したもので、本発明の範囲を限定するものではない。

【0045】サーバ205、205'、205''、…は、WWWサーバ、WWWプロキシ・サーバ、匿名FTPサーバ、ユーザの入力等の「情報源」から、クライアントに提供する情報を入手して、URLと組にして保持する。サーバ205、205'、205''、…は上記情報の入手を、上記情報源のそれぞれに定期的に（例えば1時間に一度）接続して新規情報を得る、上記情報源から新規情報の発生を通知して貰う、等の方法で行なう。なお、サーバ205、205'、205''、…がクライアントに提供する情報を入手する情報源および方法は、例として示したもので、本発明の範囲を限定するものではない。また、本実施例のサーバ205、205'、205''、…が持つ情報にはURLが付与されており、かつサーバはURLを用いてURL内容の管理を行なうが、サーバが持つ情報を何らかの形式で指定す



ることができればよく、URLに限ることはない。このため、本発明の範囲を限定するものではない。さらに本発明の実施は、クライアントまたはサーバのオペレーティング・システム、サーバ間またはサーバ・クライアント間のネットワークの種類およびネットワークの物理層プロトコルやトランスポート層プロトコル、もしくはサーバとクライアントが単一のコンピュータ上にあるか異なるコンピュータ上にあるかには依存しない。

#### 【0046】クライアントとサーバの動作の概略

図1を用いてクライアント204とサーバ205の動作の概略を説明する。なお、サーバ205とクライアント204、他のサーバ205'、205''、…を繋いでいるネットワーク202はここには図示していない。また、図1中の矢印は主要なデータの流れである。

【0047】サーバ205は依頼処理部101、プリフェッチ部102、サーバ間メッセージ処理部103の各処理部分と、ホームページキャッシュ104、アクセス戦略表105、頻度DB106、ホームページグループ表107の各データ構造を持つ。

【0048】サーバ205のホームページキャッシュ104は、サーバ205が動作する計算機203の持つ不揮発性記憶装置（磁気ハードディスク、光ディスク等。以下「二次記憶」と記すことがある）または揮発性記憶装置（メインメモリ、キャッシュメモリ等。以下「主記憶」と記すことがある）またはその両方に格納される。なお、ホームページキャッシュ104は不揮発性記憶装置に格納される場合でも揮発性記憶装置に格納される場合でも、一個所に格納される必要はない。例えば、サーバ205が動作する計算機203が3つの磁気ハードディスクを持っていれば、ホームページキャッシュ104をこれら3つの磁気ハードディスクの一部または全部に分割して格納しても差し支えない。むしろ、複数の磁気ハードディスクを用いることは、磁気ハードディスクへのアクセスの並列性を増し、もってホームページキャッシュ104の読み書きを高速化するために、必須ではないが望ましい。また当該サーバ205は情報のキャッシュをすることが重要な役割であるため、当該サーバ205には大容量のホームページキャッシュ104を構成するのに十分な記憶装置が付属していることが、必須ではないが望ましい。

【0049】クライアント204が、本システムから第1のURLに対応する情報である第1のURL内容を取得する際には、本システムの任意の第1のサーバ205に、第1のURLを含む依頼120を送信することによって、第1のURL内容の取得を依頼する。

【0050】依頼120は依頼処理部101が受け取り、ホームページキャッシュ104に依頼されたURL内容があるかどうかを検索する（130）。もしあれば、当該ホームページキャッシュ104をアクセスして、第1のURL内容をクライアント204に返答12

1で送信する。

【0051】アクセス戦略表105は、本システムの他のサーバ205'、205''、…がどのURL内容を保持しているかの記録、及び第1のサーバ205から行なった過去の通信の速度（例えばスループットやレイテンシ）の記録を保持している。第1のURL内容が第1のサーバ205のホームページキャッシュ104内になければ、第1のサーバ205はアクセス戦略表105を参照して依頼を転送すべき第2のサーバ205'、205''、…を決定する（131）。そして第2のサーバ205'、205''、…および第1のURL内容の情報源のうち1つ以上に、依頼122により依頼120を転送する。第1のサーバ205は、依頼を転送した第2のサーバ205'、205''、…または情報源の少なくとも1つから第1のURL内容が返答123によって得られると、得られた第1のURL内容をクライアント204に返答121によって送る。なお、上記依頼120、返答121、依頼122、返答123はいずれも、ネットワーク202を通じて行なわれてもかまわないし、計算機203、203'、203''、…の持つ通信機能を用いて行なわれてもかまわない。

【0052】第1のサーバ205は返答123によって第1のURL内容が得られると、第1のサーバ205のホームページキャッシュ104に第1のURL内容を追加する（133）。この場合には、当該第1のURL内容が追加されたことを伝えるホストURLメッセージ124が本システムの他のサーバ205'、205''、…の一部または全部に送信される。またホームページキャッシュ104に第1のURL内容を追加する際、別の第2のURL内容がホームページキャッシュ104から削除される場合があるが、この場合には、当該第2のURL内容が削除されたことを伝えるホストURLメッセージ124が本システムの他のサーバ205'、205''、…の一部または全部に送信される。

【0053】頻度DB106は、後に述べるように、どのURL内容がいつ参照されたかという頻度情報を記録している。ホームページグループ表107は、連続して依頼されたことが多かった複数のURL内容の組を記録している。第1のサーバ205のプリフェッチ部102は返答121の後、依頼120による頻度情報の変化を頻度DB106に反映する（135）。頻度情報の変化は、頻度DB106の更新内容を伝えるホスト頻度メッセージ125によって本システムの他のサーバ205'、205''、…の一部または全部に送信される。さらに、プリフェッチ部102は頻度DB106の内容に基づいてホームページグループ表107を更新する（136）。

【0054】次に、将来依頼されやすいと考えられるURL内容を頻度DB106およびホームページグループ表107を用いて決定する（137および138）。当

該URL内容が1個以上あれば、それらは依頼処理部101が上記依頼120と同様に受信する。

【0055】また、サーバ間メッセージ処理部103は、本システムの他のサーバ205'、205''、…からホストURLメッセージ124'を受信するとアクセス戦略表105を書き換える(140)。また、本システムの他のサーバ205'、205''、…からホスト頻度メッセージ125'を受信すると頻度DB106を書き換える(141)。

【0056】以上がクライアント204とサーバ205の動作の概略である。以下ではこれらをさらに詳しく説明する。

【0057】サーバの内部動作の詳細

図3に、サーバ205の内部構成を示す。依頼処理部101は、依頼入力部301、キャッシュ表参照部302、処理戦略部303、受信部304、返答部305、キャッシュ入れ替え部310からなる。プリフェッチ部102は、リンク統計更新部306、グループ更新部307、プリフェッチ戦略部308、プリフェッチスケジュール部309、タイマー311からなる(サーバ間メッセージ処理部103は図3には示していない)。

【0058】サーバ205のデータ構造は、ホームページキャッシュ104、アクセス戦略表105、頻度DB106(アクセス頻度表325、リンク頻度表326、外部アクセス頻度表327、外部リンク頻度表328からなる)、ホームページグループ表107、およびアクセス戦略表105を構成するためのキャッシュディレトリ323と通信コスト表324、およびホスト表の各データ構造を持つ(ホスト表は図示していない)。これらのデータ構造はすべて、主記憶と二次記憶の片方または両方に存在してかまわない。なお、図3において、太い矢印は主要な制御の流れを、また細い矢印は主要なデータの流れを表わす。また、本実施例では請求項における必要データはURL内容であり、請求項における通信履歴は、アクセス戦略表105と通信コスト表324に、データ保持確率はキャッシュディレトリ323に、時間毎通信履歴は頻度DB106に、参照関係情報はホームページグループ表107に、それぞれ格納される。また、サーバ間で通信されるデータの一覧はホストURLメッセージ124の形式で通信される。

【0059】サーバ205内部の処理の説明に先立って、各データ構造の概略を説明する。各データ構造の詳細な説明は、サーバ205内部の処理の説明とあわせて行なう。図4に、ホームページキャッシュ104、アクセス戦略表105、キャッシュディレトリ323、通信コスト表324、アクセス頻度表325、リンク頻度表326、外部アクセス頻度表327、外部リンク頻度表328の構成を示す。

【0060】ホームページキャッシュ104は、キャッシュの本体である。URL401と、URL401の内

容であるURL内容402を組にして保持する。この組は通常複数ある。

【0061】アクセス戦略表105は、下に述べるキャッシュディレトリ323と通信コスト表324とを、現在の使用のために組み合わせたものである。URL411と、ホスト412と、通信コスト413と、URL存在確率414との組を保持する。この組は通常複数ある。URL411はURLである。ホスト412はサーバ205または情報源が存在するホスト名、通信コスト413には、通信のスループットとレイテンシを用いる。URL存在確率414は、ホスト412で指定されるコンピュータに、URL411で指定されるホームページが存在する確率である。サイズ415は前回アクセス時のURL411のバイト数である。アクセス戦略表105は、第1のサーバ205が、依頼を転送する1つ以上の第2のサーバ205'、205''、…を選定する際に使用する。

【0062】なお、ホスト名には、シンボリックなホスト名かIPアドレスを用いる。場合によってはホスト名に、シンボリックなホスト名かIPアドレスと、TCP/IPやUDP/IPのポート番号とを組にしたものを用いる。ただしホスト名は通信相手のサーバまたはクライアントを指定し、通信路を開くことが目的であるため、この目的にかなえば、いかなる形態の表現をとっても差し支えない。特に、ここでIPアドレスやポート番号は例として出したものであり、本発明の範囲を限定するものではない。

【0063】キャッシュディレトリ323は、URL421と、ホスト422と、URL存在確率423を含む組を保持する。この組は通常複数ある。URL421はURLである。ホスト422はサーバまたは情報源が存在するホスト名、URL存在確率423は、ホスト422で指定されるコンピュータに、URL421で指定されるホームページが存在する確率である。

【0064】通信コスト表324は、ホスト431と、時刻432と、通信コスト433と、通信回数434を含む組を保持する。この組は通常複数ある。ホスト431はサーバまたは情報源が存在するホスト名、時刻432は何曜日の何時かを格納する。通信コスト433は時刻432においてホスト431に対して行なわれた過去一定期間内の通信のスループット(単位時間当りの通信可能データ量)とレイテンシ(通信が受ける遅延時間)である。通信回数434は、時刻432においてホスト431に対して行なわれた過去一定期間内の通信の回数を格納する。

【0065】アクセス頻度表325は、URL441と、アクセス頻度442と、サイズ443を含む組を保持する。この組は通常複数ある。URL441はURL、アクセス頻度442はURL441で指定されるホームページを、過去の一定時間C1内に当該サーバ20

5がアクセスした回数である。本実施例のアクセス頻度442は、過去一週間の何曜日の何時に何回アクセスしたかを格納するカウンタ群である。サイズ443は前回アクセス時のURL441のバイト数である。

【0066】リンク頻度表326は、参照元URL451と、参照先URL452と、アクセス頻度453を含む組を保持する。この組は通常複数ある。アクセス頻度453は参照元URL451から参照先URL452へのリンクを、過去の一定時間C2内に当該サーバ205がアクセスした回数である。本実施例のアクセス頻度453は、過去一週間の何曜日の何時に何回アクセスしたかを格納するカウンタ群である。

【0067】外部アクセス頻度表327は、URL461と、アクセス頻度462を含む組を保持する。この組は通常複数ある。URL461はURL、アクセス頻度462はURL461で指定されるホームページを、過去の一定時間C1内に本システムの他のサーバ205'、205''、…がアクセスした回数である。本実施例のアクセス頻度462は、過去一週間の何曜日の何時に何回アクセスしたかを格納するカウンタ群である。

【0068】外部リンク頻度表328は、参照元URL471と、参照先URL472と、アクセス頻度473を含む組を保持する。この組は通常複数ある。アクセス頻度473は参照元URL471から参照先URL472へのリンクを、過去の一定時間C2内に他のサーバ205'、205''、…がアクセスした回数である。本実施例のアクセス頻度473は、過去一週間の何曜日の何時に何回アクセスしたかを格納するカウンタ群である。

【0069】ホスト表は、ホスト名の組を保持する。ホスト表に現れるホスト名は、本システムに参加している計算機203、203'、203''、…上のサーバ205、205'、205''、…の全体である。

【0070】またサーバ205は、プリフェッチフラグ、新規URLフラグ、前回アクセスURLの変数を用いる。プリフェッチフラグは、真偽値の変数であり、現在サーバ205がプリフェッチ中か、クライアント204（または他のサーバ205'、205''、…）の依頼を処理中かを表わす。初期値は偽である。新規URLフラグは、受信したURL内容が新規にホームページキャッシュ104に付け加わったかどうかを示す真偽値の変数である。初期値は偽である。前回アクセスURLは、前回クライアントからの依頼に対して返答したURLである。初期値は「空」である。

【0071】図5に、ホームページグループ表107、ホストURLメッセージ124、ホスト頻度メッセージ125の構成を示す。

【0072】ホームページグループ表107は、先頭URL481とURLグループ482を含む組を保持する。この組は通常複数ある。各々の組は、連続してアク

セスされることが多いURLの集合である。先頭URL481は1つのURL、URLグループ482は1つ以上のURL483、URL483'、URL483''、…である。

【0073】ホストURLメッセージ124は、ホスト501、URL502、フラグ503を含む組を内蔵したデータである。1つのホストURLメッセージ124にこの組が2つ以上内蔵されて構わない。ホスト501はホスト名、URL502は1つのURL、フラグ503は真偽値である。ホストURLメッセージ124は、ホスト501で指定されるサーバがURL502で指定されるURLを持っていること（フラグ503が真の場合）、または持っていないこと（フラグ503が偽の場合）を他のサーバに伝達する目的で用いられる。

【0074】ホスト頻度メッセージ125は、ホスト511、参照元URL512、参照先URL513、頻度514を含む組を内蔵したデータである。ホスト511はホスト名、参照元URL512は1つのURL、参照先URL513は1つのURLまたは「空」、頻度514はカウンタである。ホスト頻度メッセージ125は、参照元URL512で指定されるURLまたは参照元URL512と参照先URL513の間のリンクを、過去の一定時間内にホスト511で指定されるサーバ205がアクセスした回数である頻度514を、他のサーバ205'、205''、…に伝える目的で用いられる。なお、上記C1、C2の値はいずれも、サーバ205のコンパイル時または起動時または起動中に変更可能であって差し支えない。また、アクセス頻度442、アクセス頻度453、アクセス頻度462、アクセス頻度473、頻度514を過去一週間の何曜日の何時に何回アクセスしたかを格納するとしたのは、一例として示したものに過ぎず、本発明の範囲を限定するものではない。また、アクセス頻度表325、リンク頻度表326、外部アクセス頻度表327、外部リンク頻度表328の目的は、どのURLやどのリンクがどのくらいアクセスされたかを記録して利用することにあるので、この目的にかなえば、他の情報を記録することも考えられる。例えば、アクセス頻度442、アクセス頻度453、アクセス頻度462、アクセス頻度473、頻度514と併用して、または代りとして、最終アクセス時刻を記録しておき、LRU（least recently used）法によってどのURLやどのリンクがどのくらいアクセスされたかを見積もることも考えられる。

【0075】以下、図1、図3、図6、および図7を用いて、第1のサーバ205がクライアント204または他のサーバ205'、205''、…からURL内容の取得を依頼された際の、第1のサーバ205の処理の流れを説明する。第1のサーバ205は依頼入力部301でクライアント204または他のサーバの依頼を待ち、受け付ける（601、350）。当該クライアント20

4または他のサーバの依頼の中で、取得すべきURL内容はURLで表現される。なおクライアント204または他のサーバからの依頼に含まれていたURLを以後「処理対象URL」と呼ぶ。第1のサーバ205において、処理対象URLはキャッシュ表参照部302に渡される(351)。この際、プリフェッチフラグには「偽」を格納する。

【0076】キャッシュ表参照部302は、ホームページキャッシュ104内で処理対象URLに等しいURL401を持つ組を検索する(602、380)。もしそのような組が存在すれば(603)、返答部305に制御を移す(363)。一方そのような組が存在しなければ(604)、処理戦略部303に処理を移す(352)。

【0077】処理戦略部303は、アクセス戦略表105を用いて(381)、依頼を転送すべき1つ以上の第2のサーバ205'、205''、…または情報源を選定する。まず、アクセス戦略表105内で処理対象URLに等しいURL411を持つ組を検索し(605)、そのような組が1つ以上存在するかどうかを調べる(606)。存在する場合(607)、後述の手順でこれら1つ以上の組の順位づけおよび選択を行ない(608)、当該順位づけおよび選択の結果選ばれた1つまたは複数の組のそれぞれに対して、算出した優先度の高い順にホスト412で指定されるサーバや情報源に依頼を転送する(610、365)。一方そのような組が1つも存在しない場合(609)、処理対象URLの情報源に依頼を転送する(610、365)。なお、上記どちらの場合でも、依頼を転送したら受信部304に処理を移す(353)。依頼を転送したサーバ205'、205''、…や情報源を、以後「依頼転送先」と呼ぶ。

【0078】順位づけおよび選択には、さまざまなアルゴリズムを使用するが、ネットワーク通信の効率面からの主要な指標はレイテンシとスループットと通信サイズである。また、依頼を転送した先に処理対象URLに対応する情報が存在する確率も重要となる。本実施例では具体的には以下に示す順位づけと選択を行なう。なお、下で用いるD1、D2、D3、D4、D5、D6はいずれも定数であるが、これらの値はいずれもサーバ205のコンパイル時または起動時または起動中に変更可能であって差し支えない。また、ここで述べる順位づけを、サーバ205が保持する他の情報と合わせてもちいても何ら差し支えない。例えば、他の情報として、システム管理者がサーバ205に提供する情報、IPアドレスの類似性、ルーティングのトポロジーやホップ数、通信回線の物理構成などを使うことが考えられる。

【0079】順序づけと選択の手順は以下の通りである。上記アクセス戦略表105内で処理対象URLに等しいURL411を持つ組から、通信コスト413に含まれるスループットが定数D1以下であるものを除外す

る。通信コスト413に含まれるレイテンシが定数D2以上であるものを除外する。URL存在確率414が定数D3以下であるものを除外する。各組の優先度(大きい方が高い)をレイテンシとスループットと通信サイズとURLの存在確率から、 $(URL存在確率414 / (通信コスト413のレイテンシ + サイズ415 / 通信コスト413のスループット))$ の式を用いて算出し、上位D4の組を選択する。次に、処理対象URLに含まれる情報源のコンピュータがこの時点で選択に含まれていなければ、当該情報源を最も優先度が低くして選択に加える。以上が本実施例における順位づけおよび選択の手順である。

【0080】なお、上記で複数のホスト412にほぼ同時に依頼を転送すると、複数の依頼転送先からほぼ同時に返答が送られて、第1のサーバ205付近のネットワークが混雑する恐れがある。一方、複数のホスト412に順々に依頼を転送し、その都度返答を待っていると、結果的に取得すべきURL内容の取得が遅くなる恐れがある。これらの問題を軽減するために、以下の通信プロトコルを用いることができる。図8を用いて説明する。この図では、サーバ205と依頼転送先との間の通信の様子を時間に沿って上から下に描いており、図の(a)と(b)は、2つの異なる場面を示している。

【0081】図8の(a)で示すように依頼を転送する際に、上記優先度の上位D5個の依頼転送先(以後優先依頼転送先(801、801'、…)と呼ぶ)には、依頼を変更なしで転送する(803)。この依頼をうけた当該優先依頼転送先のそれぞれは、処理対象URLに対応するURL内容を保持していればすぐ第1のサーバ205への返答を開始する(804)。依頼を変更なしで転送することは、処理対象URLがあればすぐに返答を開始せよ、という要求を意味する。

【0082】上記優先依頼転送先はそれぞれ、処理対象URLに対応するURL内容があれば当該URL内容を当該第1のサーバ205に返答し、なければ「処理対象URLは存在しない」という返答を返す(805)。

【0083】一方第1のサーバ205は、上記優先依頼転送先以外の依頼転送先(以後非優先依頼転送先(802、802'、…)と呼ぶ)には、処理対象URLの「返答待機依頼」(806)を送付する。返答待機依頼は、処理対象URLが存在して、かつ後述する返答開始依頼が到着したら、返答を開始せよ、という意味をもつ。この際、非優先依頼転送先は、処理対象URLに対応するURL内容があれば待機し、なければ「処理対象URLは存在しない」という返答を返す。

【0084】第1のサーバ205は図8の(b)で示すように、上記優先依頼転送先のすべてから「処理対象URLは存在しない」(805')という返答が返ってきた場合(807)に、まだ「処理対象URLは存在しない」という返答をしていない非優先依頼先の1つ以上に

対して、「返答開始依頼」(808)を送る。当該依頼が非優先依頼先に届き、かつ当該非優先依頼先に処理対象URLに対応するURL内容があれば、当該非優先依頼先は当該第1のサーバ205に対して当該URL内容の返答を開始する(809)。なお、待機中の非優先依頼転送先のおのおのは、返答待機依頼から一定時間D6以内に返答開始依頼が到着しなければ、待機を中止して次の依頼に備える。

【0085】このプロトコルは、処理対象URLに対応するURL内容が優先依頼転送先の1つ以上に存在した場合に上記第1のサーバ205付近のネットワークの混雑を緩和する。なぜなら、依頼転送先を2つの集団に分けたことによって、同時に起こる返答の最大数を減らしているためである。また、処理対象URLに対応するURL内容が優先依頼転送先のどれにも存在しなかった場合にも、処理対象URLに対応するURL内容の取得が遅くなるのを防ぐ。なぜなら、第1のサーバ205は返答待機依頼に対する非優先依頼転送先からの返答を待たずに、返答開始依頼を流すことが可能であるからである。

【0086】なお、上述の本実施例における順位づけと選択の手順によれば、優先依頼転送先または非優先依頼転送先のいずれかに、処理対象URLの情報源が含まれているので、依頼転送先のどこからも処理対象URLに対応するURL内容が転送されない場合は非常に稀であることに注意されたい。万一、処理対象URLに対応するURL内容が転送されなかった場合には、本実施例では処理対象URLに対応するURL内容の受信を断念する。なお本実施例の変形例として、この場合にさらに別の依頼転送先を選択して依頼を転送することが考えられる。また、別の変形例として、この場合に処理対象URLに対応するURL内容の受信を一定時間経過後に再度試みる方法が考えられる。

【0087】また本実施例におけるサーバ205は、当該サーバ205の設置者が利便性を享受させようとするクライアント204、204'、204''、…以外に、他のサーバへの通信を行なう。他のサーバとの通信は、サーバ205のキャッシュの能力を発揮するために必要不可欠であるが、この通信によって元来の目的としたクライアント204、204'、204''、…へのサービスが低下する恐れをはらんでいる。この問題を解決するためにサーバ205は、他のサーバからの依頼を一定時間内に一定量に抑える。本実施例のサーバ205は過去時間T1以内にL1バイト以上のURL内容の転送が、他のサーバに行なわれた場合、以後当該T1内には他のサーバからの依頼を受け付けない。また、本発明の別の実施例としては、サーバ205が過去時間T1以内にL2回数以上のURL内容の転送が、他のサーバに行なわれた場合、以後当該T1内には他のサーバからの依頼を受け付けない。なおここで、T1、L1、L2は

コンパイル時または起動時または起動中に変更可能であって差し支えない。

【0088】受信部304では、処理対象URLに対応するURL内容が、1つ以上の依頼転送先から返答されてくるのを待ち、当該URL内容の受信を行ない(366)、通信コスト表324の更新を行なう。

【0089】処理対象URLに対応する返答があるか、一定時間T2が経過するのを待つ。(611)。ここで返答があったかを調べ(612)、返答があった場合には(613)、返答に含まれるURL内容は、ホームページキャッシュ104内の新規な組として格納される(614、382)。この際、処理対象URLが当該新規な組のURL401となり、当該受信されたURL内容がURL内容402となる。なお、処理対象URLに対応するURL内容の受信後(受信中でもかまわない)にホームページキャッシュ104の総量が定数C3を越えたかを調べる(615)。越えていた場合に限り後述するキャッシュ入れ替え部310に制御が移り、キャッシュ入れ替えの処理が行なわれる(616と617、図3の355と356)。そして下に述べるように通信コスト表324の更新を行ない(618)、返答部305に制御を移す(354)。一方、処理対象URLに対応するURL内容が受信されなかった場合(619)、すなわち、依頼転送先のすべてから「処理対象URLは存在しない」という返答を得た場合には、ホームページキャッシュ104には何も操作を行わず返答部305に制御を移す(354)。なおここで、T2、C3はサーバ205のコンパイル時または起動時または起動中に変更可能であって差し支えない。

【0090】通信コスト表324の更新は、図9の手順で行なう。処理対象URLに対応するURL内容が、第1の依頼転送先から受信され始めた(910)時点で、通信コスト表324の第1の依頼転送先に対応する組を更新する。これは以下の手順で行なう。

【0091】(1)ホスト431が第1の依頼転送先に等しく、かつ時刻432が現在の時刻にあう組を検索する(911)。そして、通信コスト433のレイテンシを再計算するとともに、通信回数434を1増加する(912)。すなわち、911で該当する組が見つからなかった場合には、通信コスト表324に新たな組を作り、ホスト431を第1の依頼転送先で、時刻432を現在の時刻で、通信回数434を1で初期化する。そして、通信コスト433のレイテンシを今回のレイテンシ(すなわち当該第1の依頼転送先が優先依頼転送先ならば依頼を転送してから現在までの時間、また非優先依頼転送先ならば、転送開始依頼を送信してから現在までの時間)で初期化する。また911で該当する組が見つかった場合には、当該組の通信コスト433のレイテンシを、今回のレイテンシと、通信コスト433および通信回数434の値を用いて再計算するとともに、通信回数

434に1を加える。

【0092】(2)次に、当該第1の依頼転送先からの受信の終了時刻ENDTIMEを予測して計算する(913)。これは以下の方法で行なう。アクセス戦略表105内で処理対象URLに等しいURL411を持ち、かつホスト412が当該第1の依頼転送先に等しい組を検索する。そのような組の通信コスト413とサイズ415とを用いて、ENDTIMEを(現在の時刻+(サイズ415/通信コスト413のスループット))とする。

【0093】(3)次に、処理対象URLに対応するURL内容が第2の依頼転送先から受信され始めるか、第1の依頼転送先からの受信が終了するかのいずれかを待つ(914)。

【0094】(4)第1の依頼転送先からの受信が終了したかを調べる(915)。受信が終了しないうちに第2の依頼転送先から受信され始めた場合(916)、上記第1の依頼転送先の場合と同様の手順で、通信コスト表324の第2の依頼転送先に対応する組を更新するとともに(917)、上記第1の依頼転送先からの受信の終了時刻ENDTIMEと同様の手順で、第2の依頼転送先からの受信の終了時刻ENDTIME2を計算する(918)。そして、もしENDTIMEがENDTIME2+定数C4よりも大きいかを調べ(919)、大きければ(920)、上記第1の依頼転送先からの受信を中止し(921)、第2の転送依頼先を第1の依頼転送先に置き換え(922)、ENDTIME2をENDTIMEとし(923)、914にもどる。なおここで、C4はサーバ205のコンパイル時または起動時または起動中に変更可能であって差し支えない。また、もしENDTIMEがENDTIME2+定数C4以下であれば(924)、第2の依頼転送先からの受信を中止する(925)。

【0095】一方915において、第1の依頼転送先からの受信が終了した場合(926)には、通信コスト表324の第1の依頼転送先に対応する組の通信コスト433のスループットを、今回のスループット(すなわち、今回受信したURL内容のサイズを、(当該第1の転送依頼先が優先依頼転送先ならば)依頼を転送してから現在までの時間または(非優先転送依頼先ならば)返答開始依頼を送信してから現在までの時間で割ったもの)と、通信コスト433および通信回数434の値を用いて再計算する(927)。そして新規URLフラグを真を格納する(928)。以上が、通信コスト表324の更新手順である。

【0096】返答部305は以下の手順で動作する。プリフェッチフラグが「真」かどうかを調べる(620)。真ならば、プリフェッチ戦略部308に制御を移す(621、364)。プリフェッチフラグが「偽」であった場合(622)、処理対象URLに対応するUR

L内容が受信されたかどうか調べる(623)。受信されなかった場合(624)には、上記依頼を受けたクライアント204または他のサーバに「処理対象URLは存在しない」を返答(367)した後(625)、依頼入力部301に制御を移して新たなクライアント204または他のサーバからの依頼を待つ(626)。一方処理対象URLに対応するURL内容が受信された場合(627)、そのURL内容を上記依頼を受けたクライアント204または他のサーバに返答する(628、367)。

【0097】ここで、新規URLフラグが真かどうかを調べる(629)。新規URLフラグが真なら(630)、ホストURLメッセージ124の送信を行なったあと(631)、リンク統計更新部306に制御を移す(632、357)。新規URLフラグが偽なら(633)、すぐにリンク統計更新部306に制御を移す(357)。

【0098】新規URLフラグである場合、ホームページキャッシュ104に1つ以上の組が追加されたことを意味するので、ホストURLメッセージ124を行なう。まず新規URLフラグに偽を格納してから、1つの組を持つホストURLメッセージ124を新たに作成し、当該組のホスト501に第1のサーバ205のホスト名を、URL502に処理対象URLを、フラグ503に真をそれぞれ格納する。そして、当該組を持つホストURLメッセージ124をホスト表に含まれる他のサーバ205'、205''、…のうち1つ以上に送信する。例えば、ホスト表中の組のうち、第1のサーバ205を含む組を検索する。この組から第1のサーバ205のホスト名を除いた1つ以上のホスト名のそれぞれに対して、上記手順で作成したホスト頻度メッセージ125を送信する方法が考えられる。また別の方法ではホスト表中の組に格納されているホスト名のそれぞれに対して、当該組を持つホストURLメッセージ124を送信する方法が考えられる。また、1つの組をもつホストURLメッセージ124をすぐに他のサーバ205'、205''、…に送るのではなく、いくつかの組をまとめてから他のサーバ205'、205''、…に送ることも考えられる。最初の方法では、比較的近いサーバ205'、205''、…に通信範囲を限定することができ、また最後の方法では通信回数を削減することができる。

【0099】リンク統計更新部306では、アクセス頻度表325とリンク頻度表326を更新することにより、処理対象URLのアクセス頻度と、処理対象URLに至る特定のリンクがアクセスされる頻度を記録する(386)。以下にこの手順を述べる。

【0100】ステップ701で、アクセス頻度表325中の組のうち、URL441が処理対象URLに等しい組を検索する。このような組がなければ、新たな組をア



クセス頻度表325に作り、当該新たな組のURL441を処理対象URLで、アクセス頻度442のカウント群をすべて0で初期化する。そして検索で見つかった組または当該新たな組の、アクセス頻度442のうち現在時刻に対応するカウンタを1増加させ、サイズ443に処理対象URLに対応するURL内容のバイト数を格納する。

【0101】次に、前回アクセスURLから処理対象URLへのハイパー・テキスト・リンクがあるかを調べる(702)。関連があると判定した場合(703)、リンク頻度表326を更新する(704)。リンク頻度表326の更新は具体的には、リンク頻度表326中の組のうち、参照元URL451が前回アクセスURLに等しく、かつ参照先URL452が処理対象URLに等しい組を検索する。このような組がなければ、新たな組をリンク頻度表326中に作り、当該新たな組の参照元URL451を前回アクセスURLで、参照先URL452を処理対象URLで、アクセス頻度453のカウント群をすべて0で初期化する。そして検索で見つかった組または当該新たな組の、アクセス頻度453のうち現在時刻に対応するカウンタを1増加させる。そして、前回アクセスURLに処理対象URLを格納してから、グループ更新部307に制御を移す(705、358)。

【0102】一方、関連がないと判定した場合には(706)、前回アクセスURLに処理対象URLを格納してから、プリフェッチ戦略部308に制御を移す。

【0103】グループ更新部307では、ホームページグループ表107を更新する(707、387)。連続してアクセスされることが多いURLのひとつの集合を考えると、この集合をアクセスする際にはじめにアクセスされるURLが少数あると考えられる。本実施例ではこのような少数のURLを見つけた場合にのみリンク統計更新部306からグループ更新部307に制御を移すことを、すでに述べたリンク統計更新部306の場合わけにより行なっている。

【0104】グループ更新部307は以下の手順で行なう。ホームページグループ表107中で、先頭URL481が処理対象URLに等しい組を検索する。もしこのような組があれば、この組をホームページグループ表107中から削除する。次に、ホームページグループ表107中に新たな組を作成し、先頭URL481を処理対象URLで、URLグループ482を処理対象URLから連続してアクセスされる確率が定数C5以上の1つ以上のURLで初期化する。処理対象URLから連続してアクセスされる確率は、アクセス頻度表325と外部アクセス頻度表327とリンク頻度表326と外部リンク頻度表328を用いて計算することができる。なお、C5はサーバ205のコンパイル時または起動時または起動中に変更可能であって差し支えない。次に、第1のURLから第2のURLにリンクがある場合、第1のURL

のアクセス頻度(アクセス頻度表325中URL441が第1のURLに等しい組のアクセス頻度442のカウント群の合計)がA、第1のURLから第2のURLへのリンクのアクセス頻度(リンク頻度表326中参照元URL451が第1のURLに等しく、参照先URL452が第2のURLに等しい組のアクセス頻度453のカウント群の合計)がB、第1のURLの外部アクセス頻度(外部アクセス頻度表327中URL461が第1のURLに等しい組のアクセス頻度462のカウント)がC、第1のURLから第2のURLへのリンクの外部アクセス頻度(外部リンク頻度表328中参照元URL471が第1のURLに等しく、参照先URL472が第2のURLに等しい組のアクセス頻度473のカウントの値)がDとすると、第1のURLから第2のURLが連続してアクセスされる確率は $(B/A) \times (D/C)$ と見積もる。もしAとBの値のどちらかが0の場合、第1のURLから第2のURLが連続してアクセスされる確率は0とする。またもしCとDの値のどちらかが0の場合、 $(B/A)$ を第1のURLから第2のURLが連続してアクセスされる確率とする。

【0105】また、第1のURLから第2のURLにリンクがない場合、第3、第4、…、第NのURLが存在して、第1のURLから第3のURLへのリンク、第3のURLから第4のURLへのリンク、…、第NのURLから第2のURLへのリンクがそれぞれある場合がある。この場合には、第1のURLから第3のURLが連続してアクセスされる確率、第3のURLから第4のURLが連続してアクセスされる確率、…、第NのURLから第2のURLが連続してアクセスされる確率をそれぞれ $p_2$ 、 $p_3$ 、…、 $p_N$ とすると、第1のURLから第2のURLが連続してアクセスされる確率は $p_2 \times p_3 \times \dots \times p_N$ とする。なお、このような推移閉包にかかわる計算には、例えばAho他著の文献「データ構造とアルゴリズム」(1987年、培風館情報処理シリーズ11)第6.3章に記載のダイクストラのアルゴリズム等良く知られた方法が利用できる。

【0106】以上の処理の後、プリフェッチ戦略部308に制御を移す(359)。

【0107】プリフェッチ戦略部308では、処理対象URLがアクセスされた直後にアクセスされやすい1つ以上のプリフェッチURL群に対応するURL内容を決し、当該プリフェッチURL群をクライアント204または他のサーバからの依頼なしでホームページキャッシュ104に受信する。具体的には、以下の手順で処理を行なう。

【0108】まず、この時点でクライアント204または他のサーバからの依頼が到着しているかを調べる(708)。到着していれば(709)、プリフェッチフラグに「偽」を格納し、依頼入力部301に制御を移す。

【0109】一方依頼が到着していなければ(71

0)、プリフェッチフラグの値が真かを調べる(711)。

【0110】プリフェッチフラグが「偽」ならば(712)、ホームページグループ表107で、処理対象URLを含む組を検索し(388)、このような組に含まれる処理対象URL以外の1つ以上のURLをプリフェッチURL群とする。(713)。ここでプリフェッチフラグに「真」を格納する。

【0111】一方プリフェッチフラグが「真」ならば(714)、プリフェッチ中URLをプリフェッチURL群から削除する(715)。ここでプリフェッチURL群が空になれば、プリフェッチフラグを「偽」にする。

【0112】ここでプリフェッチURL群に1つ以上のURLが含まれるかを調べる(716)。含まれるならば(717)、プリフェッチURL群から1つのプリフェッチ中URLを取り出して(718)、プリフェッチ中URLを処理対象URLとみなしてキャッシュ表参照部302に渡して(361)、すでに述べた手順でプリフェッチ中URLの受信を開始する。含まれないならば(719)、プリフェッチスケジュール部309に制御を移す(360)。なおここまでの手順からも明らかに、このプリフェッチは、クライアントのみならず他のサーバからの依頼によっても起動される。すなわち、多段に配置された複数のサーバが、一連のプリフェッチを行なうことが可能である。

【0113】プリフェッチスケジュール部309では、今後アクセスされると考えられるURLで、現在ホームページキャッシュ104にないものを、将来ネットワークがすぐ時を狙って得るための準備を行なう。プリフェッチスケジュール部309は以下の手順で行なう。

【0114】まず、ホームページグループ表107中で、先頭URL481が前回アクセスURLに等しい組を検索することによって、前回アクセスURLに対応するホームページグループ表107の組を選択する(720)。

【0115】もしこのような組が存在すれば、URLグループ482に格納されている1つ以上の第1のURLの各々について、プリフェッチする時刻をキャッシュディレクトリ323と通信コスト表324から決定する(721、384、385)。これは以下の手順で行なう。

【0116】第1のURLに対応するURL内容がホームページキャッシュ104にあれば何もしない。一方なければ、キャッシュディレクトリ323中の組で、URL421が第1のURLに等しいものを検索する。このような組がなければ何もしない。このような組が1つ以上存在する場合には、URL存在確率423が最も大きい組を1つ選択し、その組のホスト422を将来アクセスするホストとする。次に通信コスト表324中で、ホ

スト431が当該ホストに等しい第1の組を検索する。このような組がなければ、なにもしない。もしこのような第1の組が1つ以上あれば、アクセス頻度表325中で、URL441が当該第1のURLに等しい第2の組を検索する。もしこのような第2の組があれば、第1の組のうち、 $(\text{第2の組のサイズ}443 / \text{第1の組の通信コスト}433 \text{のスループット}) + \text{第1の組の通信コスト}433 \text{のレイテンシ})$ が最小な第3の組を1つ選ぶ。第2の組がなければ、 $(\text{第1の組の通信コスト}433 \text{のレイテンシ} / \text{第1の組の通信コスト}433 \text{のスループット})$ が最小な第3の組を1つ選ぶ。

【0117】この第3の組の時刻432に次に対応する時刻を「プリフェッチ時刻」と呼ぶ。例えば、時刻432が「水曜日23:40」で、現在時刻が1月3日火曜日の3:00であれば、プリフェッチ時刻は1月4日水曜日の23:40となる。

【0118】次に、ステップ721で算出した、1つ以上の第1のURLのプリフェッチ時刻をタイマー311を設定する(722、389)。これにより、タイマー311は、すでに述べた手順で第1のURLの受信を行なうよう、時刻432の時刻に第1のURLの依頼をキャッシュ表参照部302に渡す。このタイマー311設定処理が終わったあとは、依頼入力部301に制御を移す(362)。

【0119】以上がプリフェッチスケジュール部309の処理である。

【0120】キャッシュ入れ替え

キャッシュ入れ替え部310では、ホームページキャッシュ104の部分的な削除を行なう(383)。ホームページキャッシュ104中の組の一部または全部のそれぞれについて、以下の「キャッシュ価値」を計算する。

【0121】ある第1のURLのキャッシュ価値は、アクセス戦略表105とアクセス頻度表325を用いて計算する。アクセス戦略表105内で当該第1のURLに等しいURL411を持つ1つ以上の第1の組を検索する。もしこのような組が複数あれば、それぞれの組の $(\text{通信コスト}413 \text{のレイテンシ} + (\text{サイズ}415 / \text{通信コスト}413 \text{のスループット}))$ が最小の組を1つ選択し、第1の組とする。さらに、アクセス頻度表325中で、URL441が第1のURLに等しい第2の組を検索する。

【0122】第1の組の $(\text{サイズ}415 / \text{通信コスト}413 \text{のスループット}) + \text{通信コスト}413 \text{のレイテンシ})$ がA、第1のURLのアクセス頻度(すなわち第2の組のアクセス頻度442のカウンタ群の合計)がBである場合、第1のURLのキャッシュ価値を $(A \times B) / \text{第1の組のサイズ}415$ とする。万一第1の組または第2の組がなければ、第1のURLのキャッシュ価値は0とする。

【0123】上記ホームページキャッシュ104中の組

の一部または全部のそれぞれについて「キャッシュ価値」を計算したら、この「キャッシュ価値」が最も小さい組から順にホームページキャッシュ104から削除する。ホームページキャッシュ104からの削除は、ホームページキャッシュ104の総量が定数C3以下になるまで行なう。

【0124】ホームページキャッシュ104から1つ以上の組が削除された場合、次に述べる手順でホストURLメッセージ124を送信する。ホストURLメッセージ124の送信は以下の手順で行なう。ホストURLメッセージ124を新たに作成し、上記ホームページキャッシュ104から削除された組のそれぞれに対して、ホストURLメッセージ124に新たな組を作成する。新たな組には、ホスト501に第1のサーバ205のホスト名、URL502に削除された組のURL401、フラグ503に偽を格納する。そして作成したホストURLメッセージ124をホスト表に含まれる他のサーバ205'、205''、…のうち1つ以上に送信する。送信相手の決定方法は、既にのべたホームページキャッシュ104に1つ以上の組が追加された場合のホストURLメッセージ124と同じである。

#### 【0125】タイマー処理

サーバ205は上記の一連の処理の他に、タイマー311によって一定時間ごとに行なわれる処理を行なう。なお以下に述べるタイマーによる各処理は、主フローとの競合を避ける目的で、依頼入力部301でクライアント204または他のサーバの依頼を待っている場合にのみ行なう。

【0126】タイマー311は、定期的にアクセス戦略表105をキャッシュディレクトリ323と通信コスト表324とアクセス頻度表325から再構成する。これは主に、通信コスト表324の時刻432は何曜日の何時における各ホストへの通信コストを保持するのに対して、アクセス戦略表105は現在の各ホストへの通信コストを保持するためである。本実施例では時刻432が1時間毎の情報を保持するため、アクセス頻度表325の再構成は1時間に一度行なう。アクセス戦略表105、キャッシュディレクトリ323、通信コスト表324、アクセス頻度表325の内容から、再構成の手順は明らかである。

【0127】タイマー311は、定期的にアクセス頻度表325とリンク頻度表326からホスト頻度メッセージ125を作成し、ホスト表に含まれる他のサーバ205'、205''、…の1つ以上に送信する。これは以下の手順で行なう。

【0128】まずホスト頻度メッセージ125を1つ作る。次にアクセス頻度表325中のすべての組に対して、当該ホスト頻度メッセージ125に新たな組を追加していく。追加する組のホスト511は第1のサーバ205のホスト名、参照元URL512はアクセス頻度表

325の組のURL441、参照先URL513は「空」、頻度514はアクセス頻度表325の組のアクセス頻度442のカウント群の合計をそれぞれ格納する。次にリンク頻度表326中のすべての組に対して、当該ホスト頻度メッセージ125に新たな組を追加していく。追加する組のホスト511は第1のサーバ205のホスト名、参照元URL512はリンク頻度表326の組の参照元URL451、参照先URL513はリンク頻度表326の組の参照先URL452、頻度514はリンク頻度表326の組のアクセス頻度453のカウント群の合計をそれぞれ格納する。次に、ホスト表中の組のうち、第1のサーバ205を含む組を検索する。この組から第1のサーバ205のホスト名を除いた1つ以上のホスト名のそれぞれに対して、上記手順で作成したホスト頻度メッセージ125を送信する。

【0129】また、タイマー311は、上記第1のサーバ205を含む組へのホスト頻度メッセージ125の送信よりも粗い頻度（例えば上記送信10回に1回の割合）で、ホスト表中の組のうち、第1のサーバ205を含まない組を検索し、この検索の結果である組からに格納されている1つ以上のホスト名のそれぞれに対して、上記手順で作成したホスト頻度メッセージ125を送信する。

【0130】また、タイマー311は、キャッシュディレクトリ323の組のすべてについて、定期的にURL存在確率423の再計算を行なう。これは以下の手順で行なう。キャッシュディレクトリ323中の組のそれぞれに対して、URL存在確率423を定数C6だけ減じたものをURL存在確率423に格納する。もし減じた結果が定数C7以下なら、URL存在確率423にC7を格納する。なおここで、C6およびC7はサーバ205のコンパイル時または起動時または起動中に変更可能であって差し支えない。

#### 【0131】メッセージ処理

第1のサーバ205のサーバ間メッセージ処理部103が他のサーバ205'、205''、…からホストURLメッセージ124を受け取ると、以下の手順でキャッシュディレクトリ323を更新する。ホストURLメッセージ124が内蔵する組のうちフラグ503が真の第1の組のそれぞれに対して、キャッシュディレクトリ323中の組で、ホスト422が第1の組のホスト501に等しくURL421が第1の組のURL502に等しい第2の組を検索する。このような第2の組があれば、当該第2の組のURL存在確率423を1に設定する。また、このような第2の組がなければ、キャッシュディレクトリ323中に新たな組を作成し、ホスト422をホスト501で、URL421をURL502で、URL存在確率423を1で初期化する。

【0132】また、ホストURLメッセージ124が内蔵する組のうちフラグ503が偽の第3の組のそれぞれ

に対して、キャッシュディレクトリ323中の組で、ホスト422が第1の組のホスト501に等しくURL421が第3の組のURL502に等しい第4の組を検索する。このような第4の組があれば、当該第4の組を削除する。

【0133】また、第1のサーバ205のサーバ間メッセージ処理部103が他のサーバ205'、205''、…からホスト頻度メッセージ125を受け取ると、以下の手順で外部アクセス頻度表327と外部リンク頻度表328とを更新する。

【0134】受け取ったホスト頻度メッセージ125が内蔵する組で参照先URL513が「空」の第1の組のそれぞれについて、外部アクセス頻度表327中で、URL461が第1の組の参照元URL512に等しい第2の組を検索する。このような第2の組が存在すれば、当該第2の組のアクセス頻度462に第1の組の頻度514を加え、再び当該第2の組のアクセス頻度462に格納する。一方このような第2の組が存在しなければ、外部アクセス頻度表327に新たな組を作成し、当該組のURL461を第1の組の参照元URL512で、アクセス頻度462を第1の組の頻度514で初期化する。

【0135】また、ホスト頻度メッセージ125が内蔵する組で参照先URL513が「空」以外の第3の組のそれぞれについて、外部リンク頻度表328中で、参照元URL471が第3の組の参照元URL512に等しく、参照先URL472が第3の組の参照先URL513に等しい第4の組を検索する。このような第4の組が存在すれば、当該第4の組のアクセス頻度473に第3の組の頻度514を加え、再び当該第4の組のアクセス頻度473に格納する。一方このような第4の組が存在しなければ、外部リンク頻度表328に新たな組を作成し、当該組の参照元URL471を第3の組の参照元URL512で、参照先URL472を第3の組の参照先URL513で、アクセス頻度473を第3の組の頻度514で初期化する。

【0136】

【発明の効果】本発明によれば以下の効果が得られる。

【0137】ユーザはクライアントを用いて任意の1つのサーバに接続して情報を依頼するだけで、複数のサーバが通信回線の不均一性と不安定性を隠蔽して情報を提供する。

【0138】サーバ間の通信回線のスループットとレイテンシが異なっても、サーバが相互に交換する情報により、各サーバは、最も高速に情報の取得が可能であると考えられるサーバを選択する。この結果、最も高速に情報の取得が可能であると考えられる通信回線を選択して利用することにより通信回線の不均一性が解決される。サーバが相互に交換する情報により、各サーバは例えば、時間帯や曜日によって回線の混雑が変化したり、

ある通信回線の増強によってルーティングのパターンが変更され、別の回線が混雑したり混雑が解消されたりすることを考慮に入れる。この結果、各サーバは最も高速に情報の取得が可能であると考えられる他のサーバを、最も高速に情報の取得が可能であると考えられる通信回線を選択して利用する。これにより通信回線の不安定性を回避できる。

【図面の簡単な説明】

【図1】本実施例のサーバの内部構成の概略を示すブロック図。

【図2】本実施例の分散コンピュータ・システムを示す図。

【図3】サーバ205の内部構成を示すブロック図。

【図4】サーバ205内で用いられるデータ構造を示す図である。

【図5】サーバ205内で用いられるデータ構造を示す図（図5の続き）。

【図6】サーバ205内の処理手順を示すフローチャート。

【図7】サーバ205内の処理手順を示すフローチャート（図6の続き）。

【図8】サーバ205、205'、205''、…が用いる混雑緩和のためのプロトコルを説明する図。

【図9】通信コスト表324の更新手順を示すフローチャート。

【符号の説明】

205 サーバ  
204 クライアント  
101 依頼処理部  
102 プリフェッチ部  
103 サーバ間メッセージ処理部  
104 ホームページキャッシュ  
105 アクセス戦略表  
106 頻度DB  
107 ホームページグループ表  
124 ホストURLメッセージ  
125 ホスト頻度メッセージ  
301 依頼入力部  
302 キャッシュ表参照部  
303 処理戦略部  
304 受信部  
305 返答部  
306 リンク統計更新部  
307 グループ更新部  
308 プリフェッチ戦略部  
309 プリフェッチスケジュール部  
310 キャッシュ入れ替え部  
323 キャッシュディレクトリ  
324 通信コスト表  
325 アクセス頻度表

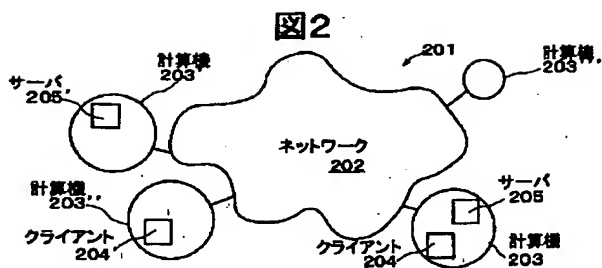
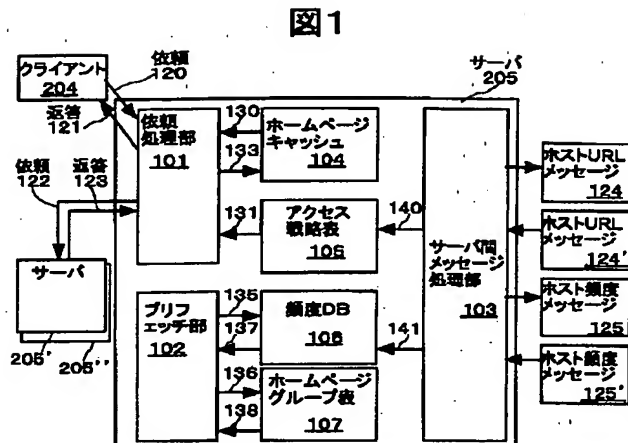
326 リンク頻度表

327 外部アクセス頻度表

328 外部リンク頻度表。

【図1】

【図2】



【図4】

【図5】

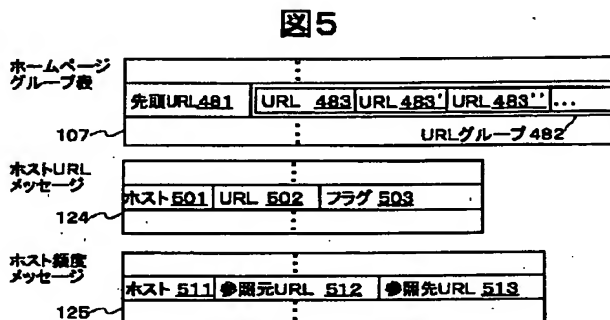
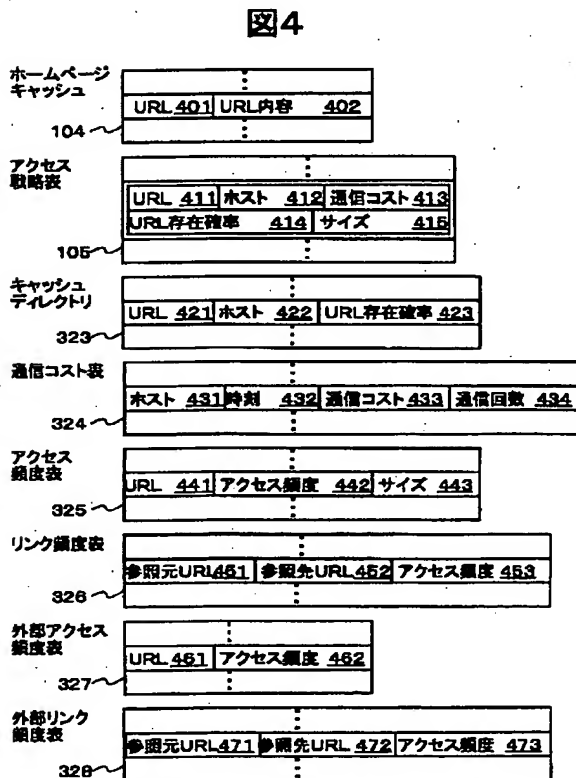


図3

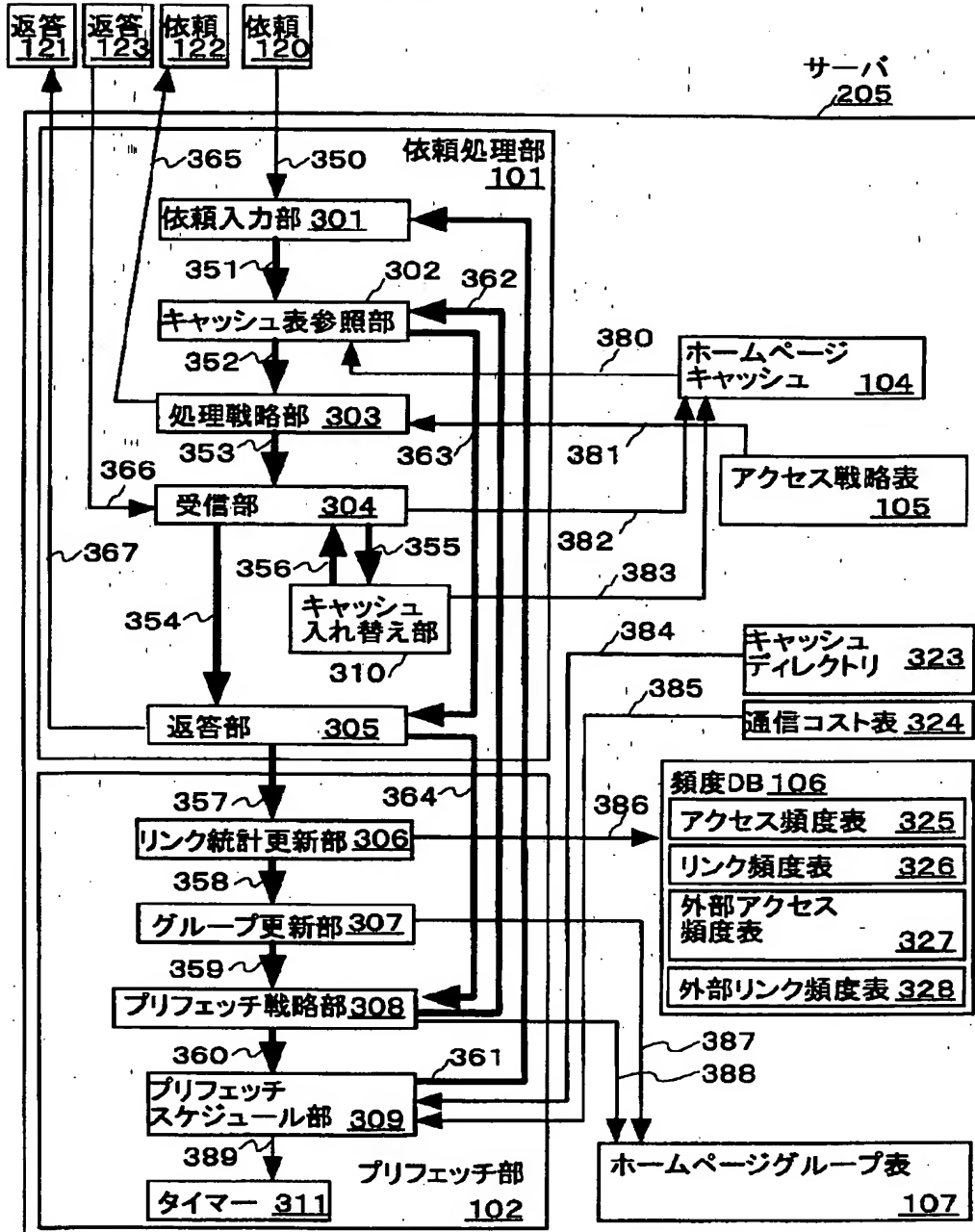




図6

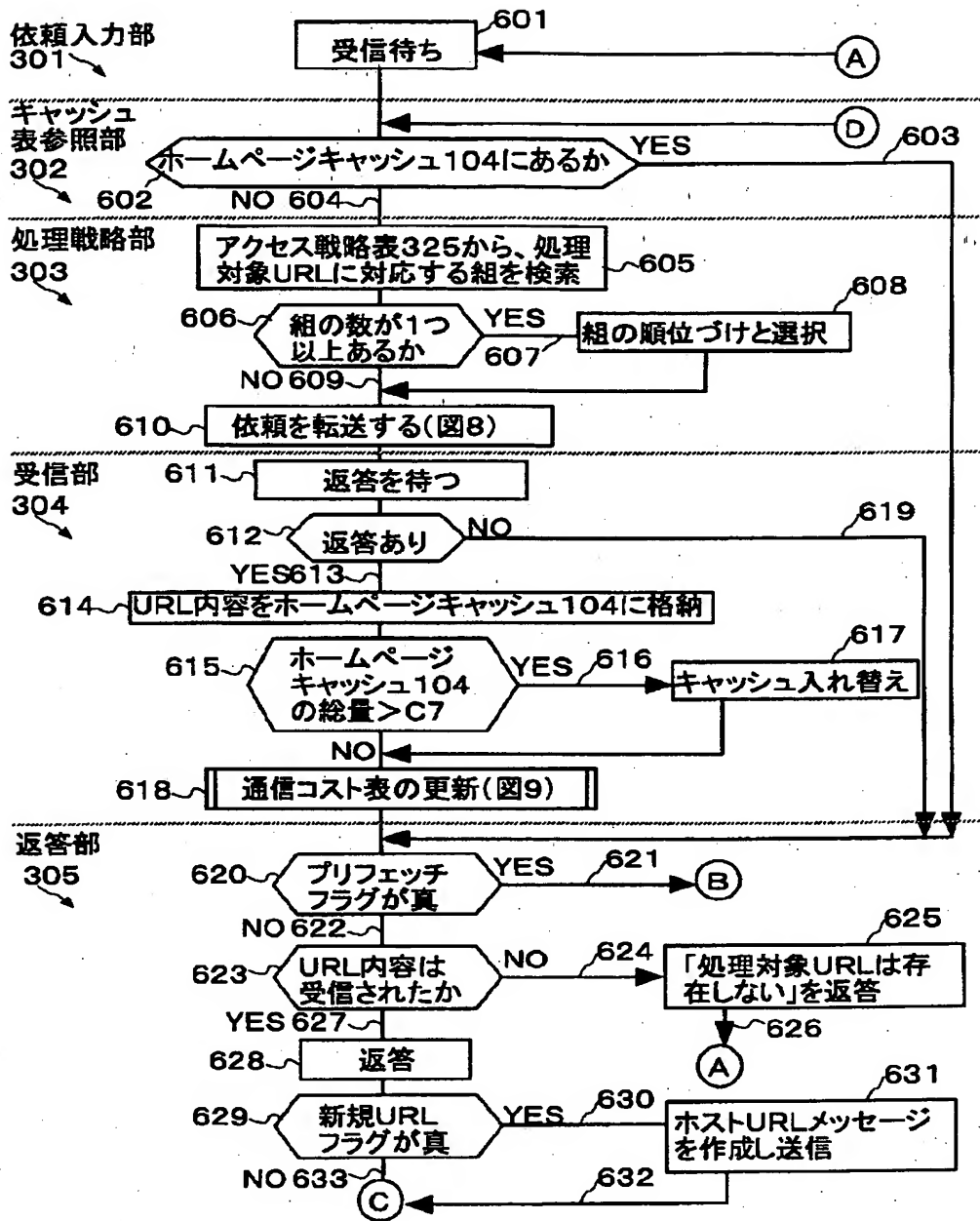
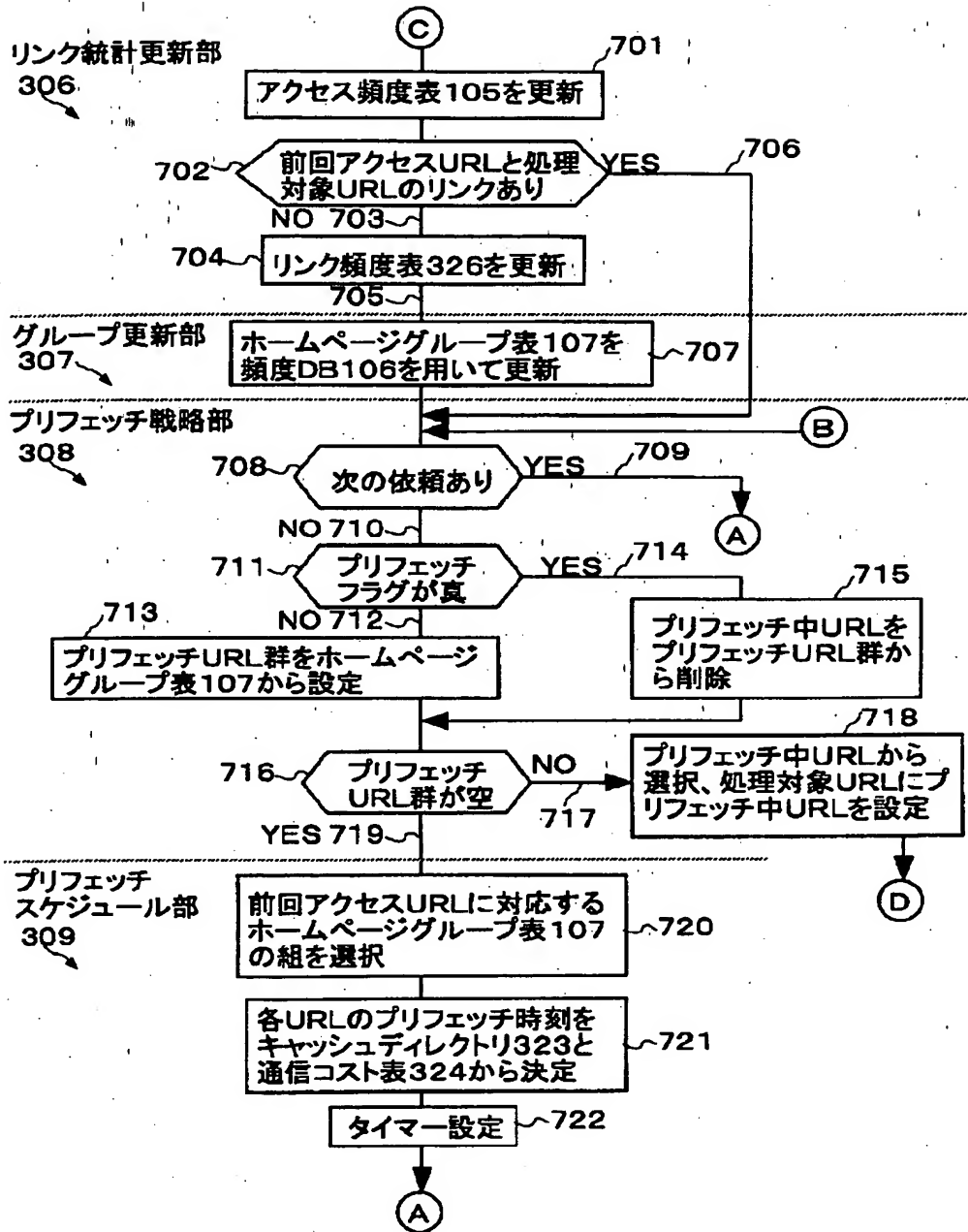


図7



【図8】

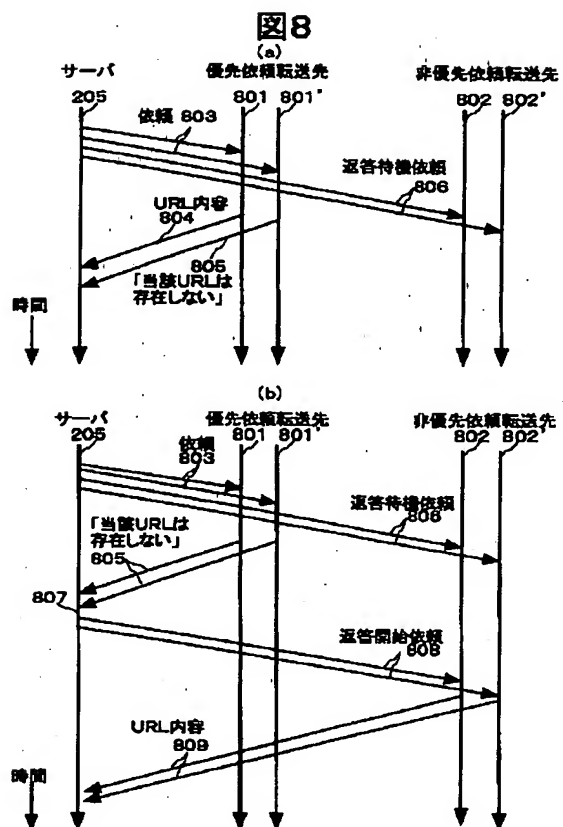
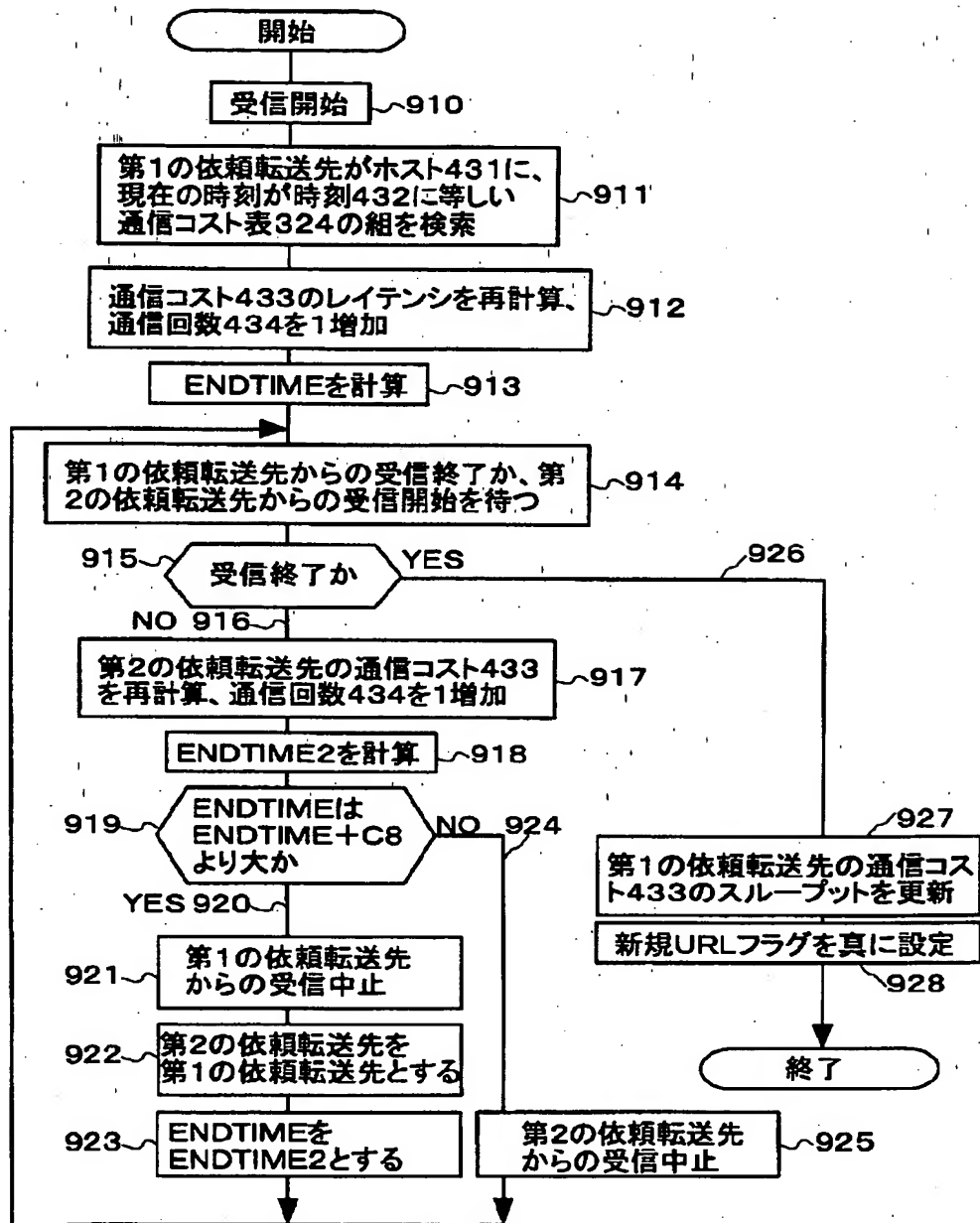


図9



フロントページの続き

(72)発明者 野田 文雄  
東京都国分寺市東恋ヶ窪一丁目280番地  
株式会社日立製作所中央研究所内

(72)発明者 関 京華  
東京都国分寺市東恋ヶ窪一丁目280番地  
株式会社日立製作所中央研究所内